SEQUENCE LISTING

<110> Juridical Foundation, Japanese Foundation For Cancer Research

<120> vector for gene therapy of malignamt melanoma, with use of virus having MSH fused protein.

<130> H11-0241J2

<160> 39

<170> PatentIn Ver. 2.0

<210> 1

<211> 166

<212> DNA

<213> Artificial Sequence

<220>

<223> DNA coding a part of adenovirus type 5 fiber, AS linker peptide and lpha-MSH.

<220>

<221> CDS

<222> (3).. (113)

<400> 1	
gg gaa ttc tcg agt tac act ttt tca tac att gcc caa gaa cca tca	47
Glu Phe Ser Ser Tyr Thr Phe Ser Tyr Ile Ala Gln Glu Pro Ser	
1 5 10 15	
gcc tcc gca tct gct tcc gcc cct gga tcc tac tcc atg gag cac ttc	95
Ala Ser Ala Ser Ala Pro Gly Ser Tyr Ser Met Glu His Phe	s •
20 25 30	
cgc tgg ggc aag ccg gtg taaagaatcg tttgtgttat gtttcaacgt	143
Arg Trp Gly Lys Pro Val	
35	
gtttattttt caattgaatt ccc	166
<210> 2	
<211> 126	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> synthetic DNA No. 924 used as temprate for PCR amplification	of DNA
sequence No. 1.	
<400> 2	
cgttgaaaca taacacaaac gattctttac accggcttgc cccagcggaa gtgctccatg	60
gagtaggatc caggggcgga agcagatgcg gaggctgatg gttcttgggc aatgtatgaa	120
aaagtg	126

```
<210> 3
```

<211> 39

<212> DNA

<213> Artificial Sequence

<220>

 $\langle 223 \rangle$ synthetic DNA No. 933 used as sense primer for PCR amplification of DNA sequence No. 1.

<400> 3

gggaattctc gagttacact ttttcataca ttgcccaag

39

<210> 4

<211> 49

<212> DNA

<213> Artificial Sequence

<220>

 $\langle 223 \rangle$ synthetic DNA No. 934 used as antisense primer for PCR amplification of DNA sequence No. 1.

<400> 4

gggaattcaa ttgaaaaata aacacgttga aacataacac aaacgattc

<210> 5

<211> 76

<212> DNA

<213 >Artificial Sequence

<220>

<223> synthetic DNA No. 1061 used as sense primer for PCR amplification of DNA coding lpha-MSH and adenovirus fiber poly A signal.

<400> 5

cgggatccta ctccatggag cacttccgct ggggcaagcc ggtgtaagtc gacaagaata 60 aagaatcgtt tgtgtt 76

<210> 6

<211> 32

<212> DNA

<213 >Artificial Sequence

<220>

<223> synthetic DNA No. 1092 used as antisense primer for PCR amplification of DNA coding α -MSH and adenovirus fiber poly A signal.

<400> 6

cggaattcat ggcgccatgt ttaatcagag gt

<210> 7

<211> 1818

<212> DNA

<213 Artificial Sequence

<220>

<223> DNA coding a modified fiber protein of pWE6.7R-F/asMSHa

<220> -

<221> CDS

<222> (1).. (1815)

<400> 7

atg aag cgc gca aga ccg tct gaa gat acc ttc aac ccc gtg tat cca 48

Met Lys Arg Ala Arg Pro Ser Glu Asp Thr Phe Asn Pro Val Tyr Pro

1 5 10 15

tat gac acg gaa acc ggt cct cca act gtg cct ttt ctt act cct ccc 96

Tyr Asp Thr Glu Thr Gly Pro Pro Thr Val Pro Phe Leu Thr Pro Pro

20 25 30

ttt gta tcc ccc aat ggg ttt caa gag agt ccc cct ggg gta ctc tct 144
Phe Val Ser Pro Asn Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
35 40 45

ttg	g cgo	cta	i teo	gaa	a cct	ct	a gtt	acc	tco	aat	ggc	atg	ctt	gcg	ctc	192
Leu	ı Arg	g Lei	ı Ser	Gli	ı Pro	Lei	ı Val	Thr	Sei	Asn	Gly	Met	Leu	Ala	Leu	
	50)				5 :	5			-	60					
aaa	ate	ggc	aac	ggc	ctc	tci	ctg	gac	gag	gcc	ggc	aac	ctt	acc	tcc	240
Lys	Met	Gly	Asn	Gly	Leu	Sei	Leu	Asp	Glu	Ala	Gly	Asn	Leu	Thr	Ser	
65					70					75		•			80	
											٠					
caa	aat	gtą	acc	act	gtg	ago	cca	cct	ctc	aaa	aaa	acc	aag	tca	aac	288
Ġln	Asn	Val	Thr	Thr	Val	Ser	Pro	Pro	Leu	Lys	Lys	Thr	Lys	Ser	Asn	
				85					90					95		
ata	aac	ctg	gaa	ata	tct	gca	ccc	ctc	aca	gtt	acc	tca	gaa	gcc	cta	336
Ile	Asn	Leu	Glu	Ile	Ser	Ala	Pro	Leu	Thr	Val	Thr	Ser	Glu	Ala	Leu	
			100					105					110			
act	gtg	gc t	gcc	gcc	gca	c c t	cta	atg	gtc	gcg	ggc	aac	aca	ctc	acc	384
Thr	Val	Ala	Ala	Ala	Ala	Pro	Leu	Me t	Val	Ala	Gly	Asn	Thr	Leu	Thr	
		115					120	-				125				
atg	caa	tca	cag	gcc	ccg	cta	acc	gtg	cac	gac	tcc	aaa	ctt	agc	att	432
Met	Gln	Ser	Gln	Ala	Pro	Leu	Thr	Val	His	Asp	Ser	Lys	Leu	Ser	Ile	
	130					135					140					
gcc	acc	caa	gga	ссс	ctc	aca	gįg	tca	gaa	gga	aag	cta	gcc	ctg	caa	480

Ala	Thr	Gln	Gly	Pro	Leu	Thr	Val	Ser	Glu	Gly	Lys	Leu	Ala	Leu	Gln	
145					150	ė				155					160	
													•			
aca	tca	ggc	ссс	ctc	acc	acc	acc	gat	agc	agt	acc	ctt	act	atc	ac t	528
Thr	Ser	Gly	Pro	Leu	Thr	Thr	Thr	Asp	Ser	Ser	Thr	Leu	Thr	Ile	Thr	
				165					170					175		
-																
gcc	tca	ccc	cct	cta	ac t	act	gcc	act	ggt	agc	ttg	ggc	att	gac	ttg	576
Ala	Ser	Pro	Pro	Leu	Thr	Thr	Ala	Thr	Gly	Ser	Leu	Gly	Ile	Asp	Leu	
			180					185					190			
aaa	gag	ccc	a t t	tat	aca	caa	aa t	gga	aaa	cta	gga	cta	aag	tac	ggg	624
Lys	Glu	Pro	Ile	Tyr	Thr	Gln	Asn	Gly	Lys	Leu	Gly	Leu	Lys	Tyr	Gly	
		195					200					205				
gct	cct	ttg	cat	gta	aca	gac	gac	cta	aac	act	ttg	acc	gta	gca	ac t	672
Ala	Pro	Leu	His	Val	Thr	Asp	Asp	Leu	Asn	Thr	Leu	Thr	Val	Ala	Thr	
	210					21.5					220					
ggt	cca	ggt	gţg	ac't	att	aat	aat	act	tcc	ttg	çaa	ac t	aaa	gtt	act	720
Gly	Pro	Gly	Val	Thr	Ile	Asn	Asn	Thr	Ser	Leu	Gln	Thr	Lys	Val	Thr	
225					230					235					240	
gga	gcc	ttġ	ggt	ttt	gat	tca	caa	ggc	aat	atg	caa	ctt	aat	gta	gca	768
					Asp											
				245					250					255		

gg	a gg	a ct	a ag	g at	t ga	t tc	t ca	a aa	c ag	a cg	c ct	t at	a ct	t ga	t gtt	816
Gl	y Gl	y Le	u Ar	g Il	e Ası	p Se	r Gl	n Asi	n Ar	g Ar	g Lei	u Ile	e Lei	ı As	p Val	
			26	0			•	26	5				270)		
agi	t ta	t cc	g tt	t ga	t gci	t caa	a aac	c caa	a cta	a aa	t cta	a aga	ı cta	ı gg	a cag	864
															y Gln	
		27					280					285				
ggc	cct	ct	t tt	t ata	a aac	: tca	gcc	cac	aac	ttg	g gat	att	aac	tac	c aac	912
															Asn	
	290					295					300					
aaa	ggc	ctt	tac	ttg	ttt	aca	gct	tca	aac	aat	tcc	aaa	aag	ctt	gag	960
															Glu	
305					310					315					320	
						÷										
gtt	aac	cta	ago	act	gcc	aag	ggg	ttg	atg	t t t	gac	gct	aca	gcc	ata	1008
				Thr												
				325					330					335		
gcc	a t t	aat	gca	gga	gat	ggg	ctt	gaa	ttt	ggt	tca	cct	aat	gca	cca	1056
				Gly												
			340					345	•				350			
			•										•			
aac	aca	aat	ccc	ctc	aaa	aca	aaa	a t t	ggc	cat	ggc	cta	gaa	t t t	ga t.	1104

As	n Th	r Ası	ı Pro	Let	ı Lys	Thi	Lys	Ile	Gly	His	Gly	/ Leu	Glu	ı Phe	e Asp	
		355	5				360)				365				
t c	a aa	c aag	g gc t	atg	gtt	cct	aaa	cta	gga	act	ggo	ctt	agt	ttt	gac	1152
Se	r Ası	ı Lys	Ala	Met	Val	Pro	Lys	Leu	Gly	Thr	Gly	Leu	Ser	Phe	Asp	
	370)				375					380	ſ				
ag	c aca	ıggt	gcc	att	aca	gta	gga	aac	aaa	aat	aat	gat	aag	cta	act	1200
Se	r Thr	Gly	Ala	Ile	Thr	Val	Gly	Asn	Lys	Asn	Asn	Asp	Lys	Leu	Thr	
38	5				390					395					400	
t t g	g tgg	acc	aca	cca	gc t	cca	tct	cct	aac	tgt	aga	cta	aat	gca	gag	1248
Lei	ı Trp	Thr	Thr	Pro	Ala	Pro	Ser	Pro	Asn	Cys	Arg	Leu	Asn	Ala	Glu	
				405					410					415		
aaa	gat	gc t	aaa	ctc	act	ttg	gtc	tta	aca	aaa	tgt	ggc	agt	caa	ata	1296
Lys	Asp	Ala	Lys	Leu	Thr	Leu	Val	Leu	Thr	Lys	Cys	Gly	Ser	Gln	Ile	
			420					425					430			
ctt	gc t	aca	gtt	tca	gtt	ttg	gct	gtt	aaa	ggc	agt	ttg	gct	cca	ata	1344
Leu	Ala	Thr	Val	Ser	Val	Leu	Ala	Val	Lys	Gly	Ser	Leu	Ala	Pro	Ile	
		435					440		٠			445				
tct	gga	aca	gtt	caa	agt	gc t	cat	ctt	att	ata	aga	ttt	gac	gaa	aat	1392
	Gly															
	450					455					460					

													•			
gg	a gtg	g cta	cta	aao	aat	tcc	tto	cte	, gac	cca	gaa	tat	t gg	aac	t t t	1440
Gl	/ Val	Leu	Leu	Asr	a Asn	Ser	Phe	e Leu	Asp	Pro	Glu	Tyr	Trp	Asn	Phe	
46	5				470	I				475	ı				480	
aga	ı aat	gga	gat	ctt	act	gaa	ggo	aca	gcc	tat	aca	aac	gct	gtt	gga	1488
Arg	s Asn	Gly	Asp	Leu	Thr	Glu	Gly	Thr	Ala	Tyr	Thr	Asn	Ala	Val	Gly	
				485					490					495		
ttt	atg	cct	aac	cta	tca	gct	tat	cca	aaa	tct	cac	ggt	aaa	act	gcc	1536
		Pro														
			500					505					510			
aaa	agt	aac	a t t	gtc	agt	caa	gtt	tac	tta	aac	gga	gac	aaa	act	aaa	1584
Lys	Ser	Asn	Ile	Val	Ser	Gln	Val	Tyr	Leu	Asn	Gly	Asp	Lys	Thr	Lys	
		515					520					525				
cct	gta	aca	cta	acc	att	aca	cta	aac	ggt	aca	cag	gaa	aca	gga	gac	1632
Pro	Val	Thr	Leu	Thr	I l [·] e	Thr	Leu	Asn	Gly	Thr	Gln	Glu	Thr	Gly	Asp	
	530					535					540					
aca	ac t	cca	agt	gca	tac	tct	atg	tca	ttt	tca	tgg	gac	t gg	tct	ggc	1680
		Pro														
545					550					555					560	
			•												_	
cac	aac	tac	att	aat	gaa	ata	ttt	gcc	acc	tcg	agt	tac	act	ttt	tca	1728

His Asn Tyr Ile Asn Glu Ile Phe Ala Thr Ser Ser Tyr Thr Phe Ser 565 570 575

tac att gcc caa gaa cca tca gcc tcc gca tct gct tcc gcc cct gga 1776

Tyr Ile Ala Gln Glu Pro Ser Ala Ser Ala Ser Ala Ser Ala Pro Gly

580 585 590

tcc tac tcc atg gag cac ttc cgc tgg ggc aag ccg gtg taa 1818

Ser Tyr Ser Met Glu His Phe Arg Trp Gly Lys Pro Val

595 600 605

<210> 8

<211> 40

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic DNA No. 1037 used as sense primer for PCR amplification of DNA coding human MSH receptor residue 1-154.

<400> 8

gggaattcac catggctgtg cagggatccc agagaagact

40

<210> 9

<211> 33 <212> DNA <213> Artificial Sequence <220> <223 synthetic DNA No. 1038 used as antisense primer for PCR amplification of DNA coding human MSH receptor residue 150-317. <400> 9 gggaattcac caggagcatg tcagcacctc ctt <210> 10 <211> 27

33

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic DNA No. 1039 used as sense primer for PCR amplification of DNA coding human MSH receptor residue 150-317.

<400> 10

ctgcggtacc acagcatcgt gaccctg

27

<210> 11

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic DNA No. 1040 used as antisense primer for PCR amplification of DNA coding human MSH receptor residue 1-154.

<400> 11

gctgtggtac cgcagtgcgt agaagat

27

<210> 12

<211> 107

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic DNA No. 1075 used as sense primer for PCR amplification of DNA coding β -MSH and adenovirus fiber poly A signal.

<400> 12

cgcggatccg ccgagaagaa ggacgagggc ccctacagga tggagcactt ccgctggggc 60
agcccgccca aggactaagt cgacaagaat aaagaatcgt ttgtgtt 107

<210> 13

<211> 1848

<212> DNA

<213> Artificial Sequence

<220>

<223> DNA coding a modified fiber protein of pWE6.7R-F/asMSHb

<220>

<221> CDS

<222> (1).. (1845)

<400> 13

atg aag cgc gca aga ccg tct gaa gat acc ttc aac ccc gtg tat cca 48

Met Lys Arg Ala Arg Pro Ser Glu Asp Thr Phe Asn Pro Val Tyr Pro

1 5 10 15

tat gac acg gaa acc ggt cct cca act gtg cct ttt ctt act cct ccc 96

Tyr Asp Thr Glu Thr Gly Pro Pro Thr Val Pro Phe Leu Thr Pro Pro

20 25 30

ttt gta tcc ccc aat ggg ttt caa gag agt ccc cct ggg gta ctc tct 144
Phe Val Ser Pro Asn Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
35 40 45

ttg cgc cta tcc gaa cct cta gtt acc tcc aat ggc atg ctt gcg ctc 192 Leu Arg Leu Ser Glu Pro Leu Val Thr Ser Asn Gly Met Leu Ala Leu

aaa	atg	ggc	aac	ggc	ctc	tct	ctg	gac	gag	gcc	ggc	aac	ctt	acc	tcc	240
Lys	Met	Gly	Asn	Gly	Leu	Ser	Leu	Asp	Glu	Ala	Gly	Asn	Leu	Thr	Ser	
65					70					75					80	
							•									
caa	aat	gta	acc	ac t	gtg	agc	cca	cct	ctc	aaa	aaa	acc	aag	tca	aac	288
Gln	Asn	Val	Thr	Thr	Val	Ser	Pro	Pro	Leu	Lys	Lys	Thr	Lys	Ser	Asn	
				85					90					95		
ata	aac	ctg	gaa	a t a	tct	gca	ccc	ctc	aca	gtt	acc	tca	gaa	gcc	cta	336
Ile	Asn	Leu	Glu	Ile	Ser	Ala	Pro	Leu	Thr	Val	Thr	Ser	Glu	Ala	Leu	
			100					105					110			
														-		
act	gtg	gct	gcc	gcc	gca	cct	cta	atg	gtc	gcg	ggc	aac	aca	ctc	acc	384
Thr	Val	Ala	Ala	Ala	Ala	Pro	Leu	Met	Val	Ala	Gly	Asn	Thr	Leu	Thr	
		115				·	120					125				
atg	caa	tca	cag	gcc	ccg	cta	acc	gtg	cac	gac	tcc	aaa	ctt	agc	at t	432
Met	Gln	Ser	Gln	Ala	Pro	Leu	Thr	Val	His	Asp	Ser	Lys	Leu	Ser	Ile	
	130					135					140					
gcc	acc	caa	gga	ссс	ctc	aca	gtg	tca	gaa	gga	aag	cta	gcc	ctg	caa	480
Ala	Thr	Gln	Gly	Pro	Leu	Thr	Val	Ser	Glu	Gly	Lys	Leu	Ala	Leu	Gln	

aca	. tca	ggc	ccc	cto	acc	acc	acc	gat	ago	agt	acc	ctt	act	atc	act	528
Thr	Ser	Gly	Pro	Leu	Thr	Thr	Thr	Asp	Ser	Ser	Thr	Leu	Thr	Ile	Thr	
				165	i				170	ı			•	175		
gcc	tca	ссс	cct	cta	act	act	gcc	act	ggt	agc	ttg	ggc	att	gac	ttg	576
Ala	Ser	Pro	Pro	Leu	Thr	Thr	Ala	Thr	Gly	Ser	Leu	Gly	Ile	Asp	Leu	
			180					185				£	190			
aaa	gag	ссс	att	tat	aca	caa	aat	gga	aaa	cta	gga	cta	aag	tac	ggg	624
Lys	Glu	Pro	Ιŀe	Tyr	Thr	Gln	Asn	Gly	Lys	Leu	Gly	Leu	Lys	Tyr	Gly	
		195					200					205				
gc t	cct	ttg	cat	gta	aca	gac	gac	cta	aac	act	ttg	acc	gta	gca	act	672
Ala	Pro	Leu	His	Val	Thr	Asp	Asp	Leu	Asn	Thr	Leu	Thr	Val	Ala	Thr	
	210					215					220					
ggt	cca	ggt	gtg	ac t	att	aat	aat	act	tcc	ttg	caa	ac t	aaa	gtt	ac t	720
Gly	Pro	Gly	Val	Thr	Ile	Asn	Asn	Thr	Ser	Leu	Gln	Thr	Lys	Val	Thr	
225					230					235					240	
gga	gcc	ttg	ggt	ttt	gat	tca	caa	ggc	aat	atg	caa	ctt	aat	gta	gca	768
Gly	Ala	Leu	Gly	Phe	Asp	Ser	Gln	Gly	Asn	Met	Gln	Leu	Asn	Val	Ala	
				245					250					255		
gga	gga	cta	agg	a t t	gat	tct	caa	aac	aga	cgc	ctt	ata	ctt	gat	gtt	816
Gly	Gly	Leu	Arg	lle	Asp	Ser	Gln	Asn	Arg	Arg	Leu	Ile	Leu	Asp	Val	

								•								
agt	tat	ccg	ttt	gat	gc t	caa	aac	caa	cta	aat	cta	aga	ċta	gga	cag	864
Ser	Tyr	Pro	Phe	Asp	Ala	Gln	Asn	Gln	Leu	Asn	Leu	Arg	Leu	Gly	Gln	
	٠	275					280					285				
ggc	cct	ctt	ttt	ata	aac	tca	gcc	cac	aac	ttg	gat	a t t	aac	tac	aac	912
Gly	Pro	Leu	Phe	Ile	Asn	Ser	Ala	His	Asn	Leu	Asp	Ile	Asn	Tyr	Asn	
	290					295					300					
aaa	ggc	ctt	tac	ttg	t t t	aca	gc t	tca	aac	aat	tcc	aaa	aag	ctt	gag	960
Lys	Gly	Leu	Tyr	Leu	Phe	Thr	Ala	Ser	Asn	Asn	Ser	Lys	Lys	Leu	Glu	
305					310					315					320	
				•							·					
gtt	aac	cta	agc	ac t	gcc	aag	ggg	ttg	atg	ttt	gac	gc t	aca	gcc	a t a	1008
Val	Asn	Leu	Ser	Thr	Ala	Lys	Gly	Leu	Met	Phe	Asp	Ala	Thr	Ala	Ile	
				325					330					335		
gcc	att	aat	gca	gga	gat	ggg	ctt	gaa	ttt	ggt	tca	c c t	aat	gca	cca	1056
Ala	Ile	Asn	Ala	Gly	Asp	Gly	Leu	Glu	Phe	Gly	Ser	Pro	Asn	Ala	Pro	

265

270

350

260

340

aac aca aat ccc ctc aaa aca aaa att ggc cat ggc cta gaa ttt gat 1104 Asn Thr Asn Pro Leu Lys Thr Lys Ile Gly His Gly Leu Glu Phe Asp 355 360 365

	tca	aac	aag	gc t	atg	gtt	cct	aaa	cta	gga	ac t	ggc	ctt	agt	t t t	gac	1152
	Ser	Asn.	Lys	Ala	Met	Val	Pro	Lys	Leu	Gly	Thr	Gly	Leu	Ser	Phe	Asp	
		370					375					380					
	agc	aca	ggt	gcc	att	aca	gta	gga	aac	aaa	aat	aat	gat	aag	cta	act	1200
,	Ser	Thr	Gly	Ala	Ile	Thr	Val	Gly	Asn	Lys	Asn	Asn	Asp	Lys	Leu	Thr	
	385					390					395					400	
	ttg	tgg	acc	aca	cca	gct	cca	t c t	cct	aac	tgt	aga	cta	aat	gca	gag	1248
	Leu	Trp	Thr	Thr	Pro	Ala	Pro	Ser	Pro	Asn	Cys	Arg	Leu	Asn	Ala	Glu	
					405					410					415		
							•										
	aaa	gat	gct	aaa	ctc	act	ttg	gtc	tta	aca	aaa	tgt	ggc	agt	caa	ata	1296
	Lys	Asp	Ala	Lys	Leu	Thr	Leu	Val	Leu	Thr	Lys	Cys	Gly	Ser	Gln	Ile	
				420					425					430			
4	ctt	gc t	aca	gtt	tca	gtt	ttg	gc t	gtt	aaa	ggc	agt	ttg	gc t	cca	ata	1344
.]	Leu	Ala	Thr	Val	Ser	Val	Leu	Ala	Val	Lys	Gly	Ser	Leu	Ala	Pro	Ile	
		٠.	435					440					445				
	tct	gga	aca	gtt	caa	agt	gc t	cat	ctt	att	ata	aga	ttt	gac	gaa	aat	1392
	Ser	Gly	Thr	Val	Gln	Ser	Ala	His	Leu	Ile	Ile	Arg	Phe	Asp	Glu	Asn	
		450					455					460					
	gga	gtg	cta	cta	aac	aat	tcc	ttc	ctg	gac	cca	gaa	tat	tgg	aac	ttt	1440
												Glu					
	-									-			-	•			

465					470					475					480	
aga	aat	gga	gat	ctt	act	gaa	ggc	aca	gcc	tat	aca	aac	gc t	gtt	gga	1488
Arg	Asn	Gly	Asp	Leu	Thr	Glu	Gly	Thr	Ala	Tyr	Thr	Asn	Ala	Val	Gly	
				485					490					495		
t t t	atg	cct	aac	cta	tca	gc t	tat	cca	aaa	tct	cac	ggt	aaa	ac t	gcc	1536
Phe	Met	Pro	Asn	Leu	Ser	Ala	Tyr	Pro	Lys	Ser	His	Gly	Lys	Thr	Ala	
			500					505					510			
aaa	agt	aac	att	gtc	agt	caa	gtt	tac	tta	aac	gga	gac	aaa	ac t	aaa	1584
Lys	Ser	Asn	Ile	Val	Ser	Gln	Val	Tyr	Leu	Asn	Gly	Asp	Lys	Thr	Lys	
		515					520					525				
cct	gta	aca	cta	acc	a t t	aca	cta	aac	ggt	aca	cag	gaa	aca	gga	gac	1632
Pro	Val	Thr	Leu	Thr	Ile	Thr	Leu	Asn	Gly	Thr	Gln	Glu	Thr	Gly	Asp	
	530					535					540					
aca	act	cca	agt	gca	tac	tct	atg	tca	ttt	tca	tgg	gac	t gg	t c t	ggc	1680
Thr	Thr	Pro	Ser	Ala	Tyr	Ser	Met	Ser	Phe	Ser	Trp	Asp	Trp	Ser	Gly	
545					550					555					560	
cac	aac	tac	att	aat	gaa	ata	ttt	gcc	acc	tcg	agt	tac	ac t	t t t	tca	1728
His	Asn	Tyr	Ile	Asn	Glu	Ile	Phe	Ala	Thr	Ser	Ser	Tyr	Thr	Phe	Ser	
				565					570					575		

tac att gcc caa gaa cca tca gcc tcc gca tct gct tcc gcc cct gga 1776

Tyr Ile Ala Gln Glu Pro Ser Ala Ser Ala Ser Ala Ser Ala Pro Gly

580 585 590

tcc gcc gag aag aag gac gag ggc ccc tac agg atg gag cac ttc cgc 1824
Ser Ala Glu Lys Lys Asp Glu Gly Pro Tyr Arg Met Glu His Phe Arg
595 600 605

tgg ggc agc ccg ccc aag gac taa 1848

Trp Gly Ser Pro Pro Lys Asp
610 615

<210> 14

<211> 61

<212> DNA

<213 Artificial Sequence

<220>

<223> synthetic DNA No. 1060 used as antisense primer for PCR amplification of DNA coding a part of adenovirus type 5 fiber and GS linker peptide.

<400> 14

cgggatccag atccagaacc actaccactt ccagaacctt cttgggcaat gtatgaaaaa 60

<210> 15

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic DNA No. 1098 used as antisense primer for PCR amplification of DNA coding a part of adenovirus type 5 fiber and GS linker peptide.

<400> 15

cgtgtggatc cgctgccaga accactacca cttccagaac c

41

<210> 16

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> synthetic DNA No. 931 used as sense primer for PCR amplification of DNA coding a part of adenovirus type 5 fiber and GS linker peptide.

<400> 16

ggcctttact tgtttacagc

<210> 17 <211> 1818 <212> DNA <213 >Artificial Sequence <220> <223> DNA coding a modified fiber protein of pWE6.7R-F/gsMSHa <220> <221> CDS <222> (1).. (1815) <400> 17 atg aag cgc gca aga ccg tct gaa gat acc ttc aac ccc gtg tat cca 48 Met Lys Arg Ala Arg Pro Ser Glu Asp Thr Phe Asn Pro Val Tyr Pro 1 5 10 15 tat gac acg gaa acc ggt cct cca act gtg cct ttt ctt act cct ccc 96 Tyr Asp Thr Glu Thr Gly Pro Pro Thr Val Pro Phe Leu Thr Pro Pro 20 25 30 ttt gta tcc ccc aat ggg ttt caa gag agt ccc cct ggg gta ctc tct 144 Phe Val Ser Pro Asn Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser 35

tig cgc cta tcc gaa cct cta gtt acc tcc aat ggc atg ctt gcg ctc 192

45

Leu	ı Arg	g Lei	u Se	r Glu	ı Pro	Leu	Val	Thr	Ser	Asr	Gly	Met	Let	ı Ala	Leu	
	5()				5 5	•				60)				
aaa	atg	g ggo	c aac	c ggo	ctc	tct	ctg	gac	gag	gcc	ggo	: aac	cti	acc	tcc	240
Lys	Met	Gly	/ Asr	ı Gly	/ Leu	Ser	Leu	Asp	Glu	Ala	Gly	Asn	Leu	ı Thr	Ser	
65					70					75					80	
caa	aat	gta	acc	act	gtg	agc	cca	cct	ctc	aaa	aaa	acc	aag	g tca	aac	288
Gln	Asn	Val	Thr	Thr	Val	Ser	Pro	Pro	Leu	Lys	Lys	Thr	Lys	Ser	Asn	
				85					90			•		95		
			•													
ata	aac	ctg	gaa	ata	tct	gca	ccc	ctc	aca	gtt	acc	tca	gaa	gcc	cta	336
Ile	Asn	Leu	Glu	Ile	Ser	Ala	Pro	Leu	Thr	Val	Thr	Ser	Glu	Ala	Leu	
			100					105					110			
act	gţg	gc t	gcc	gcc	gca	cct	cta	atg	gtc	gcg	ggc	aac	aca	ctc	acc	384
Thr	Val	Ala	Ala	Ala	Ala	Pro	Leu	Met	Val	Ala	Gly	Asn	Thr	Leu	Thr	
		115					120					125				
atg	caa	tca	cag	gcc	ccg	cta	acc	gtg	cac	gac	tcc	aaa	ctt	agc	at t	432
Met	Gln	Ser	Gln	Ala	Pro	Leu	Thr	Val	His	Asp	Ser	Lys	Leu	Ser	Ile	
	130					135					140					
gcc	acc	caa	gga	ссс	ctc	aca	gtg	tca	gaa	gga	aag	cta	gcc	ctg	caa	480
Ala	Thr	Gln	Gly	Pro	Leu	Thr	Val	Ser	Glu	Gly	Lys	Leu	Ala	Leu	Gln	
145					150					155					160	•

a	ca t	ca	ggo	cc	C C1	ic ad	cc a	cc a	cc.	gat	ag	c ag	gt ac	cc c	c t t	ac	t at	c	act	528
Th	ır S	er	Gly	Pr	o Le	eu Th	ır Ti	nr T	hr .	Asp	Se.	r Se	r Th	ır [Leu	.Th	r Il	e	Thr	
					16						170						17			
					·															
gc	c to	ca	ссс	СС	t ct	a ac	t ac	t go	cc a	act	gg	t ag	c tt	g g	gc	a t i	t ga	c	ttg	576
							r Th													*
				180						85						190		•		
aa	a ga	g	ссс	a t t	ta	t ac	a ca	a aa	ıt g	ga	aaa	. cta	a gg	ас	t a	ลลย	t ta	c	ggg	624
							r Gl													024
			195					20			-				05	2,5		•	O I y	
														Σ,	• •					
gci	cc	t t	tg	cat	gta	aca	ı ga	c ga	СС	t a	aac	act	t t s	z ac	: c	o ta	or:		a c t	672
							Asp													672
	210						215						220		••		7110		1111	
ggt	cca	ıg	gt	gtg	act	att	aat	aaı	t ac	c t	tcc	ttg	саа	ac	t :	ลลล	a t t	9	no t	720
							Asn													720
225						230						235	OII	111	1 1	<i>-</i> y S	Yaı			
												200						۷	40	
gga	gcc	t	tg g	ggt	ttt	gat	tca	caa		rc a	na f	ain	000	o t			-4-			7.00
							Ser													768
					245	пор	501	OIII	01			Mei	GIII	rei	u A			A	l a	
					~ 10					7	50						255			
gga	gga	c t	a 2	σσ ·	a t t	an t	t a t	0.5.												
gga	00"	υı	u a	55 ⁽	шіі	gaı	ıcı	caa	aa	c a	ga (cgc	ctt	ata	С	tt	ga t	g	t t	816

Gly Gly Leu Arg Ile Asp Ser Gln Asn Arg Arg Leu Ile Leu Asp Val	
260 265 270	
210	
agt tat ccg tit gat gct caa aac caa ata aat	
agt tat ccg tit gat gct caa aac caa cta aat cta aga cta gga cag 8 Ser Tyr Pro Phe Asp Ala Cla Asp Cla	64
Ser Tyr Pro Phe Asp Ala Gln Asn Gln Leu Asn Leu Arg Leu Gly Gln 275	
280 285	
ggc cct ctt ttt ata aac tca gcc cac aac ttg gat att aac tac aac 91	2
Gly Pro Leu Phe Ile Asn Ser Ala His Asn Leu Asp Ile Asn Tyr Asn	
290 295 300	
aaa ggc ctt tac ttg ttt aca gct tca aac aat tcc aaa aag ctt gag 960	•
Lys Gly Leu Tyr Leu Phe Thr Ala Ser Asn Asn Ser Lys Lys Leu Glu	J
310	
315 320	
gtt aac cta agc act gcc aag ggg tto -to-to-	
gtt aac cta agc act gcc aag ggg ttg atg ttt gac gct aca gcc ata 100 Val Asn Leu Ser Thr Ala Iva Clark	8
Val Asn Leu Ser Thr Ala Lys Gly Leu Met Phe Asp Ala Thr Ala Ile	
330 335	
gcc att aat gca gga gat ggg cit gaa tit ggt ica cct aat gca cca 1056	;
Ala Ile Asn Ala Gly Asp Gly Leu Glu Phe Gly Ser Pro Asn Ala Pro	
340 345 350	
aac aca aat ccc ctc aaa aca aaa att ggc cat ggc cta gaa ttt gat 1104	
Asn Thr Asn Pro Leu Lys Thr Lys Ile Gly His Gly Leu Glu Phe Asp	
355	
360 365	

t c	a aa	ic aa	ig g	ct a	tg gt	t cc	t aa	ıa ct	a gg	a ac	t gg	c ct	t ag	t tt	t gac	1152
															e Asp	
	37		•			37					38					
ag	c ac	a gg	t go	c at	t ac	a gta	a gg	a aa	c aa	a aa	t aa	t ga:	t aag	g cta	a act	1200
															Thr	
38					39					39					400	
tts	g tg	g ac	c ac	а сс	a gc	t cca	tc	t cc	t aac	tg:	t aga	ı cta	ı aat	gca	gag	1248
					o Ala											
				40					410					415		
			•													
aaa	gat	gc	taa	a ct	c act	ttg	gto	c t t a	aca	aaa	tgt	ggc	agt	caa	a t a	1296
Lys	Asp	Ala	Ly:	s Lei	ı Thr	Leu	Val	Leu	Thr	Lys	Cys	Gly	Ser	Gln	Ile	
			420)				425					430			
ctt	gc t	aca	gtt	tca	gtt	ttg	gct	gtt	aaa	ggc	agt	ttg	gct	cca	a t a	1344
Leu	Ala	Thr	Val	Ser	Val	Leu	Ala	Val	Lys	Gly	Ser	Leu	Ala	Pro	Ile	
		435					440					445				
tct	gga	aca	gtt	caa	agt	gc t	cat	ctt	a t t	a t a	aga	ttt	gac	gaa	aat	1392
Ser	Gly	Thr	Val	Gln	Ser	Ala	His	Leu	Ile	lle	Arg	Phe	Asp	Glu	Asn	
	450					455					460					
gga	gįg	cta	cta	aac	aat	tcc	ttc	ctg	gac	cca	gaa	tat	t gg	aac	t t t	1440

	Gly	/ Val	Leu	Leu	Asn	ı Asn	Ser	Phe	Leu	Asp	Pro	Glu	Tyr	Trp	Asn	Phe	
	465	j				470					475					480	
	aga	ı aat	gga	gat	ctt	act	gaa	ggc	aca	gcc	tat	aca	aac	gct	gtt	gga	1488
	Arg	g Asn	Gly	Asp	Leu	Thr	Glu	Gly	Thr	Ala	Tyr	Thr	Asn	Ala	Val	Gly	
					485			٠		490					495		
	ttt	a t _. g	cct	aac	cta	tca	gc t	tat	cca	aaa	tct	cac	ggt	aaa	act	gcc	1536
	Phe	Met	Pro	Asn	Leu	Ser	Ala	Tyr	Pro	Lys	Ser	His	Gly	Lys	Thr	Ala	
				500					505					510			
			•														
			aac														1584
	Lys	Ser	Asn	Ile	Val	Ser	Gln	Val	Tyr	Leu	Asn	Gly	Asp	Lys	Thr	Lys	
			515					520					525				
			aca														1632
	Pro		Thr	Leu	Thr	Ile	Thr	Leu	Asn	Gly	Thr	Gln	Glu	Thr	Gly	Asp	
		530					535					540					
			cca														1680
		Thr	Pro	Ser	Ala		Ser	Met	Ser	Phe	Ser	Trp	Asp	Trp	Ser	Gly	
	545					550					555					560	
			tac														1728
]	His	Asn	Tyr	He		Glu	Ile	Phe	Ala	Thr	Ser	Ser	Туг	Thr	Phe	Ser	
					565					570					575		

tac att gcc caa gaa ggt tct gga agt ggt agt ggt tct ggc agc gga 1776 Tyr Ile Ala Gln Glu Gly Ser Gly Ser Gly Ser Gly Ser Gly 580 585 590

tcc tac tcc atg gag cac ttc cgc tgg ggc aag ccg gtg taa 1818 Ser Tyr Ser Met Glu His Phe Arg Trp Gly Lys Pro Val 595 600 605

<210> 18

<211> 1848

<212> DNA

<213> Artificial Sequence

<220>

<223> DNA coding a modified fiber protein of pWE6.7R-F/gsMSHb

<220>

<221> CDS

<222> (1).. (1845)

<400> 18

atg aag cgc gca aga ccg tct gaa gat acc ttc aac ccc gtg tat cca 48 Met Lys Arg Ala Arg Pro Ser Glu Asp Thr Phe Asn Pro Val Tyr Pro 1 5 10

28/107

t	a t	ga	a	cg	gaa	a	c g	gt c	c t	cca	act	t gt	g co	et t	tto	: t t	ac	t c	c t	ccc	96
T	уr	Asp) TI	hr	Glu	Th	r G	ly Pi	ro]	Pro	Thr	· Va	l Pr	o Pl	ne L	eu	·Th	r P	ro	Pro)
					20						25							0			
t i	tt,	gta	to	c (ссс	aa	t gg	g tt	t c	aa	gag	ag	t cc	ССС	t g	gg	gt	a ci	t c	tet	144
Ph	ie '	Val	Se	r I	Pro	As	n Gl	y Ph	e G	ln	Glu	Se	r Pr	o Pr	o G	lv	Va	1 I.e	211	Ser	111
				5						40						45		- 20		001	
													•								
t t	g	gç	c t	a t	сс	gaa	а сс	t ct	a g	t t	acc	tco	aa	t gg	c a	tg	cti	gr	φ	ctc	192
Le	u A	rg	Le	u S	er	Glu	ı Pr	o Lei	u V	a i	Thr	Ser	Ası	ı Gl	v Me	e t	Ler	ı Al	ล	Len	132
		50						5 9						. 60					u	Dea	
aaa	a a	t g	ggo	a	ac	ggc	cto	tct	: c1	g	gac	gag	gco	ggo	c aa	.C	ctt	acı	r	tee	240
Lys	s M	e t	Gly	' As	sn	Gly	Leu	Ser	Lε	u <i>i</i>	Asp	Glu	Ala	Gly	. As	n	Len	Thi	r	Ser	240
65							70						75				204	1111		80	
																				00	
caa	aa	it ;	gta	ac	с	act	gtg	agc	СС	a c	ct	ctc	aaa	aaa	ac	c a	ลลฮ	tea	נ ו	aac	288
Gln	As	n '	Va l	Th	r]	ſhr	Val	Ser	Pr	o P	ro i	Leu	Lys	Lys	Th	r J	vs	Ser	. 1	len	200
						85						90		•		_	., 0	95		1311	
																		00			
a t a	aa	СС	tg	ga	a a	t a	tct	gca	ccc	СС	tc a	ıca	git	acc	toa	ρ	าลล	gr.r	c	t a	226
Ile	As	n L	eu	Glu	u I	l e	Ser	Ala	Pro) L	eu T	`hr	Val	Thr	Ser	. c	111	<u> </u>	ī	ια	336
				100			•				05	•	-•	4	JU1		10	ma	L	cu	
																1	10				

384

act gtg gct gcc gca cct cta atg gtc gcg ggc aac aca ctc acc

A

T	hr	Val	Al	a A	la A	l a	Ala	Pro) Le	eu M	e t	Va:	l Al	a G	lу	Asr	ı Th	ır	Leu	Thr	
			11	5					12	0						125	<u>,</u>				
a	tg	caa	tc	a ca	g g	сс	ccg	cta	ac	c g	t g	cac	ga	c t	СС	aaa	ct	t	agc	att	432
																				Ιle	
		30						135							10						
gc	c a	сс	caa	a gg	a co	c c	ctc	aca	gt	gto	a	gaa	gg	a aa	ıg (cta	gc	C ·	ctg	caa	480
																				Gln	
14							50						15							160	
ac	a t	сa	ggc	CC	c ct	c a	сс	acc	acc	ga	t a	agc	agt	ac	C	ett	acı	t a	atc	act	528
) Le																
					16							170							75		
gco	: to	ca	ссс	cct	cta	a a	ct a	ac t	gcc	ac	t g	ggt	agc	t t į	g g	gc	att	g	ac	ttg	576
					Lei																
				180						185							190				
aaa	ga	g (ССС	a t t	t a t	ac	a c	aa	aat	gga	a	aa	cta	gga	ı c	t a	aag	t.	ac ;	ggg	624
Lys																					
			95						200							05				-	
gc t	СС	t t	t g	cat	gta	ac	a g	ac g	gac	cta	aa	ac a	ac t	ttg	ac	CC §	gta	go	ca a	ıc t	672
Ala																					
	210						2 1							220					_	-	

gg	t cc	a gg	gt gi	gac	t at	t aa	t aa	t ac	t tc	c tt	g ca	a ac	t aa	a gt	t act	720
Gl	y Pr	o G	y Va	ıl Th	r Il	e As	n As	n Th	r Se	r Le	u Gli	n Th	r.Ly	s Va	l Thr	
22	5				23	0				23	5				240)
gg	a gc	c t t	g gg	ttt	t ga	t tc	a ca	a gg	c aai	t at	g caa	a cti	t aa	t gta	a gca	768
															Ala	
				24					250					255		
															,	
gga	ı gga	a ct	a ag	g at	t ga	t tc	t caa	a aac	: aga	ron	ctt	ata	cti	t gat	· ~++	010
														ı gal ı Asp		816
			26		,		. 011	265		, Ale	s rea	. 116			vai	
								200					270)		
agt	tat	rr c	or fift	t en	re t				- 4 -							
														gga		864
501	1 9 1			, war) Ala	. GIN			Leu	Asn	Leu	Arg	Leu	Gly	Gln	
		275)				280					285				
														tac		912
Gly		Leu	Phe	He	Asn	Ser	Ala	His	Asn	Leu	Asp	Ile	Asn	Туг	Asn	
	290					295					300	`				
aaa	ggc	ctt	tac	itg	ttt	aca	gct	tca	aac	aat	tcc	aaa	aag	ctt	gag	960
Lys	Gly	Leu	Tyr	Leu	Phe	Thr	Ala	Ser	Asn	Asn	Ser	Lys	Lys	Leu	Glu	
305					310					315					320	
gtt	aac	cta	agc	act	gcc	aag	ggg	ttg	atg	ttt	gac	gct	aca	gcc	ata	1008

Val	Asn	Leu	Ser	Thr	Ala	Lys	Gly	Leu	Met	Phe	Asp	Ala	Thr	Ala	Ile
				325					330					335	
•							•								

gcc	att	aat	gca	gga	gat	ggg	ctt	gaa	ttt	ggt	tca	c c t	aat	gca	cca	1056
Ala	Ile	Asn	Ala	Gly	Asp	Gly	Leu	Glu	Phe	Gly	Ser	Pro	Asn	Ala	Pro	
٠			340					345					350			

aac	aca	aat	ccc	ctc	aaa	aca	aaa	at t	ggc	cat	ggc	cta	gaa	t t t	gat	1104
Asn	Thr	Asn	Pro	Leu	Lys	Thr	Lys	Ile	Gly	His	Gly	Leu	Glu	Phe	Asp	
		355					360					365				

tca	aac	aag	gc t	atg	gtt	cct	aaa	cta	gga	act	ggc	ctt	agt	t t t	gac	1152
Ser	Asn	Lys	Ala	Met	Val	Pro	Lys	Leu	Gly	Thr	Gly	Leu	Ser	Phe	Asp	
	370					375					380					

agc	aca	ggt	gcc	att	aca	gta	gga	aac	aaa	aat	a a t	gat	aag	cta	act	1200
Ser	Thr	Gly	Ala	Ile	Thr	Val	Gly	Asn	Lys	Asn	Asn	Asp	Lys	Leu	Thr	
385					390					395					400	

ttg	tgg	acc	aca	cca	gc t	cca	tct	cct	aac	tgt	aga	cta	aat	gca	gag	1248
Leu	Trp	Thr	Thr	Pro	Ala	Pro	Ser	Pro	Asn	Cys	Arg	Leu	Asn	Ala	Glu	
				405					410					415		

aaa	gat	gc t	aaa	ctc	ac t	ttg	gtc	tta	aca	aaa	tgt	ggc	agt	caa	ata	1296
Lys	Asp	Ala	Lys	Leu	Thr	Leu	Val	Leu	Thr	Lys	Cys	Gly	Ser	Gln	Ile	
			420					425					430			

ctt	gc t	aca	gtt	tca	gtt	ttg	gct	gtt	aaa	ggc	agt	ttg	gc t	cca	ata	1344
Leu	Ala	Thr	Val	Ser	Val	Leu	Ala	Val	Lys	Gly	Ser	Leu	Ala	Pro	Ile	
		435					440					445				
tct	gga	aca	gtt	caa	agt	gc t	cat	ctt	at t	ata	aga	t t t	gac	gaa	aat	1392
Ser	Gly	Thr	Val	Gln	Ser	Ala	His	Leu	Ile	Ile	Arg	Phe	Asp	Glu	Asn	
	450					455					460					
					•											
gga	gtg	cta	cta	aac	aat	tcc	ttc	ctg	gac	cca	gaa	tat	t gg	aac	ttt	1440
Gly	Val	Leu	Leu	Asn	Asn	Ser	Phe	Leu	Asp	Pro	Glu	Tyr	Trp	Asn	Phe	
465					470					475					480	
-								÷								
aga	aat	gga	gat	ctt	ac t	gaa	ggc	aca	gcc	tat	aca	aac	gc t	gtt	gga	1488
Arg	Asn	Gly	Asp	Leu	Thr	Glu	Gly	Thr	Ala	Tyr	Thr	Asn	Ala	Val	Gly	
				485					490					495		
ttt	atg	c c t	aac	cta	tca	gc t	tat	cca	aaa	tct	cac	ggt	aaa	ac t	gcc	1536
Phe	Met	Pro	Asn	Leu	Ser	Ala	Tyr	Pro	Lys	Ser	His	Gly	Lys	Thr	Ala	
			500					505					510			
aaa	agt	aac	at t	gtc	agt	caa	gtt	tac	t t a	aac	gga	gac	aaa	ac t	aaa	1584
Lys	Ser	Asn	Ile	Val	Ser	Gln	Val	Tyr	Leu	Asn	Gly	Asp	Lys	Thr	Lys	
		515					520					525				
c c t	gta	аса	cta	acc	att	aca	cta	2 2 C	aa t	a.c.a	cag	നമാ	202	aas	an o	1620

Pro Val Thr Leu Thr Ile Thr Leu Asn Gly Thr Gln Glu Thr Gly Asp 530 535 540

aca	act	cca	agt	gca	tac	tct	atg	tca	t t t	tca	t gg	gac	tgg	tct	ggc	1680
Thr	Thr	Pro	Ser	Ala	Tyr	Ser	Met	Ser	Phe	Ser	Trp	Asp	Trp	Ser	Gly	
545					550					555					560	

cac aac tac att aat gaa ata ttt gcc acc tcg agt tac act ttt tca 1728 His Asn Tyr Ile Asn Glu Ile Phe Ala Thr Ser Ser Tyr Thr Phe Ser 565 570 575

tac att gcc caa gaa ggt tct gga agt ggt agt ggt tct ggc agc gga 1776

Tyr Ile Ala Gln Glu Gly Ser Gly Ser Gly Ser Gly Ser Gly Ser Gly 580 585 590

tcc gcc gag aag aag gac gag ggc ccc tac agg atg gag cac ttc cgc 1824 Ser Ala Glu Lys Lys Asp Glu Gly Pro Tyr Arg Met Glu His Phe Arg 595 600 605

tgg ggc agc ccg ccc aag gac taa

1848

Trp Gly Ser Pro Pro Lys Asp
610 615

<210> 19

<211> 128

<212> DNA

<213> Artificial Sequence

<220>

 $\langle 223 \rangle$ synthetic DNA No. 1089 used as sense primer for PCR amplification of DNA coding K21 linker peptide.

<400> 19

<210> 20

<211> 1893

<212> DNA

<213> Artificial Sequence

<220>

<223> DNA coding a modified fiber protein of pWE6. 7R-F/asK21MSHa

<220>

<221> CDS

<222> (1).. (1890)

<400> 20

	atg	g a	ag	cg	c go	ca a	ga c	cg	t c t	gaa	a ga	it a	сс	ttc	aao	C CC	c g	t g	tat	cca	. 48
	Me	L	ys	Ar	g Al	a A	rg P	ro	Ser	Glu	ı As	рТ	hr	Phe	Ası	n Pro	o Va	il (Туг	Pro	
	1						5						10						15		
1	at	ga	ıc	ace	g ga	a ac	c g	gto	cct	cca	ac	t g	tg (cct	ttt	ctt	ac	t c	ec t	ccc	96
7	yr	As	p	Thr	Gl	u Th	ır G	ly I	Pro	Pro	Th	r Va	al I	Pro	Phe	Leu	. Th	r P	ro	Pro	30
					2						2						3		10	110	
																	Ū	Ū			
t	t t	gt	a i	tcc	CC	c aa	t gg	g t	t t	caa	gag	g ag	t c	сс	cct	ggg	øt:	ас	f.c	tot	144
P	he	٧a	1 5	Ser	Pro) As	n Gl	уΡ	he	Gln	Glu	ı Se	r P	ro	Pro	Gly	Va	11.	ρ11	Sar	144
				35						40						45	, α	LL	cu	261	
																10					
t	g	cgo	С	ta	tcc	gaa	а сс	t c	ta ;	git	acc	t c	c a	at:	gge	atg	c t t	a.	·	o t o	1.00
Le	eu .	Arg	, L	eu	Ser	Glu	Pr	o Le	eu 1	Val	Thr	Se	r A:	sn (Glv	Met	Len	. δ 0	og i		192
		50							55						60	MC L	LCu	AI	la. 1	Leu	•
													•		•						
aa	a a	atg	gg	gc	aac	ggc	cto	tc	t c	tg :	gac	gag	go	oc e	7gr	aac	ctt	2.0	o f	0.0	0.40
Ly	s M	le t	G i	lу.	Asn	Gly	Leu	Se	r L	eu /	Asp	Glu	A 1	a G	llv.	Asn .	T 011	at.	- c		240
6							70				•			'5	119 1	1311	Leu	1 11	1. 3		
													'	U						80	
caa	ı a	a t	gt	a a	ıcc	act	gtg	age	СС	ca c	ec f	ctc	ลล	ລ ວ	22 1	icc a		4			0.5.5
Glr	ı A	sn	۷a	1 7	`hr	Thr	Val	Sei	r Pi	rn P	ro.	I AII	L	u a	T	hr I	iag	ıca	aa.	ac	288
						85		_ 0 /	- •	. 0 1	1 0	90	Ly.	s L	y 5 I	nr L	LYS			sn	
												30					·	95)		
ata	aa	a C	cts	g g	aa :	ata	tet	თ ი ა	CC		to -		~ 4 '			ca g					
		•		_ 0	- '	4		5 ca		i C	i C &	ıca	git	ı ac	c t	ca g	aa	gcc	c t	а	336

Ile Asn Leu Glu Ile Ser Ala Pro Leu Thr Val Thr Ser Glu Ala Leu

)0	105	110
, •	100	

act gig gct gcc gcc gca cct cta aig gic gcg ggc aac aca cic acc	384
Thr Val Ala Ala Ala Pro Leu Met Val Ala Gly Asn Thr Leu Thr	004
115 120 125	
•	
atg caa tca cag gec ccg cta acc gtg cac gac tcc aaa ctt agc att	400
Met Gln Ser Gln Ala Pro Leu Thr Val His Asp Ser Lys Leu Ser Ile	432
130	
135 140	
gcc acc caa gga ccc ctc aca gtg tca gaa gga aag cta gcc ctg caa	
Ala Thr Gln Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu Gln	480
145	
155 160	
aca tra ggr crc ete ees	
aca toa ggo coo cto acc acc gat ago agt acc ctt act atc act	528
Thr Ser Gly Pro Leu Thr Thr Asp Ser Ser Thr Leu Thr Ile Thr	
165 170 175	
gcc tca ccc cct cta act act gcc act ggt agc ttg ggc att gac ttg	576
Ala Ser Pro Pro Leu Thr Thr Ala Thr Gly Ser Leu Gly Ile Asp Leu	
180 185 190	
100	
	324
aaa gag ccc att tat aca caa aat gga aaa ota gga att	524

gct cct ttg cat gta aca	700					
gct cct ttg cat gta aca	gac gac	cta aac	act ttg	acc gta	gca act	672
Ala Pro Leu His Val Thr		Leu Asn	Thr Leu	Thr Val	Ala Thr	
210	215		220	,		
ggt cca ggt gtg act att	aat aat	act tcc	ttg caa	act aaa	gtt act	720
Gly Pro Gly Val Thr Ile	Asn Asn	Thr Ser	Leu Gln	Thr Lys	Val Thr	
225 230			235		240	
					510	
gga gcc itg ggt itt gat	tca caa	ggc aat	atg caa	ctt aat	ata ann	760
Gly Ala Leu Gly Phe Asp	Ser Gln	Glv Asn	Met Gln	Ten Ass	Val Ala	768
245		250	Mot OIII			
		200			255	
gga gga cla agg att gat	tet ean	000 000		,		
gga gga cta agg att gat	Cor Cln	aac aga (gc cii a	ata ctt	gat gtt	816
Gly Gly Leu Arg Ile Asp S 260			Arg Leu	Ile Leu	Asp Val	
200	4	265		270		
agt tot oog till mil						
agt tat ccg tit gat gct c	aa aac c	caa cta a	at cta a	iga cta g	ga cag	864
Ser Tyr Pro Phe Asp Ala G	ln Asn G	Sln Leu A	sn Leu A	rg Leu G	ly Gln	
275	280		2	85		
ggc cct ctt ttt ata aac to	ca gcc c	ac aac ti	tg gat a	tt aac t	ac aac	912
Gly Pro Leu Phe Ile Asn Se	er Ala H	is Asn Le	eu Asp I	le Asn T	vr Asn	
290 29			300		71 71011	
aaa ggc cti tac ttg ttt ac	a get te	:A AAC 22	t too ==	10 000	ı	
Lys Gly Leu Tvr Len Pho Th	r Ala Ca	r han h	r rec ga	ia aag ct	t gag	960
Lys Gly Leu Tyr Leu Phe Th	r via Se	I ASII ASI	n Ser Ly	s Lys Le	u Glu	

30	5				31	0				31	5				320	
															c ata	1008
٧a	l Asi	ı Le	u Se	r Thi	r Ala	a Lys	Gly	y Lei	ı Mei	Pho	e Asi	o Al	a Th	r Al	a Ile	
				328	5				330)				33	5	
gco	att	aa	t gca	a gga	gai	ggg	ctt	gaa	ıttt	ggt	tea	ı cc	t aa	t gca	a cca	1056
															a Pro	
			340					345					350		•	
aac	aca	aat	ccc	ctc	aaa	. aca	aaa	att	ggc	cat	ggo	cta	a gaa	ıttt	gat	1104
															Asp	1101
		355					360				-	365			ор	
tca	aac	aag	gct	atg	gtt	cct	aaa	cta	gga	act	ggc	ctt	agt	† † †	gac	1152
															Asp	1102
	370					375	_,_	200	013	1111	380	Lcu	501	1116	АЗР	
				•							300					
agc	аса	ggt	gcc	att	a c a	ata	aa.	220	000	201						
															act	1200
	1111	Gly	MId	116		vai	ыу	Asn	Lys	Asn	Asn	Asp	Lys	Leu	Thr	
385					390					395					400	
ttg	tgg	acc	aca	cca	gc t	cca	tct	cct	aac	tgt	aga	cta	aat	gca	gag	1248
Leu	Trp	Thr	Thr	Pro	Ala	Pro	Ser	Pro	Asn	Cys	Arg	Leu	Asn	Ala	Glu	
				405					410					415		

aa	a ga	t gc	t aa	a ct	c ac	t tt	ggt	c tta	a ac	a aa	a tg	t gg	c ag	t ca	a ata	1296
Ly	s As	p Al	a Ly	s Le	u Th	r Le	u Va	l Lei	ı Th	r Ly	s Cy	s Gl	y Se	r Gl	n Ile	
			42	0				428	5				430	0		
ct	gc	t ac	a gt	t tc	a gt	t tts	g gct	gtt	aa	a gg	c ag	tttg	g gc	t cc	a ata	1344
															o Ile	
		43	5				440)				445	5			
tct	gga	a aca	ı gt	t caa	a ag	t gct	cat	ctt	a t	t ata	a aga	a ttt	gao	ga	a aat	1392
															ı Asn	
	450					455					460				•	
gga	gtg	cta	cta	a aac	aat	tcc	ttc	ctg	gac	cca	a gaa	tat	t gg	aac	ttt	1440
															Phe	
465					470					475					480	
aga	aat	gga	gat	ctt	act	gaa	ggc	aca	gcc	tat	aca	aac	gct	gtt	gga	1488
Arg	Asn	Gly	Asp	Leu	Thr	Glu	Gly	Thr	Ala	Tyr	Thr	Asn	Ala	Val	Gly	
				485										495		
						,										
ttt	atg	cct	aac	c t a	tca	gct	tat	cca	aaa	tct	cac	ggt	aaa	act	gcc	1536
						Ala										
			500					505					510			
aaa	agt	aac	att	gtc	agt	caa	gtt	tac	tta	aac	gga	gac	aaa	act	aaa	1584
						Gln										- 0 0 1

515	520	525
		0.50

	ı gı	a a	a c	aac	cat	t ac	a ct	a aa	c gg	t ac	a ca	g ga	a ac	a gg	a gac	1632
Pr	o Va	.l Th	ır Le	u Th	r II	e Th	r Le	u As	n Gl	y Th	r Gl	n Gl	u Th	r Gl	y Asp	
	53	0				53	5				54	0				
			,													
ac	a ac	t cc	a ag	t gc	a ta	c tc	t at	g tc	a tti	ttca	a tg	g gao	c tg	g tc	t ggc	1680
															r Gly	
54		,			55					555					560	
cao	c aac	c ta	c at	t aa	t gaa	a a'ta	a tt	gco	acc	tce	g agt	tac	act	t tt1	tca	1728
															e Ser	
				565	5				570					575	; ;	
tac	att	gco	caa	a gaa	cca	ı tca	gcc	tcc	gca	tct	gct	tcc	gcc	cct	gga	1776
Tyr	Ile	Ala	Glr	Glu	Pro	Ser	Ala	Ser	Ala	Ser	Ala	Ser	Ala	Pro	Gly	
			580)				585					590			
					٠											
t c t	gga	tct	aag	aag	aag	aag	aag	aaa	aag	aag	aaa	aag	aag	aag	aag	1824
												Lys				
		595					600					605				
aaa	aaa	aag	aag	aag	aaa	aag	aaa	gga	tcc	tac	tcc	atg	gag	cac	ttc	1872
												Met				- + · -
	610					615					620				-	

cgc tgg ggc aag ccg gtg taa Arg Trp Gly Lys Pro Val 625 630

1893

<210> 21

<211> 1893

<212> DNA

<213> Artificial Sequence

<220>

<223> DNA coding a modified fiber protein of pWE6.7R-F/gsK21MSHa

<220>

<221> CDS

<222> (1).. (1890)

<400> 21

atg aag cgc gca aga ccg tct gaa gat acc ttc aac ccc gtg tat cca 48

Met Lys Arg Ala Arg Pro Ser Glu Asp Thr Phe Asn Pro Val Tyr Pro

1 5 10 15

tat gac acg gaa acc ggt cct cca act gtg cct ttt ctt act cct ccc 96

Tyr Asp Thr Glu Thr Gly Pro Pro Thr Val Pro Phe Leu Thr Pro Pro
20 25 30

t	t t	gta	ιto	c c	cc a	a t	ggg	g tt	t ca	aa g	ag a	ıgt	СС	c co	ct g	gg	gta	a ct	С	tct	144
															0 G						
				5						10						ļ5 ·					
t	tg (cgc	c t	a to	c g	aa (c c t	cta	ıgt	t ac	cc t	сс	aa	t gg	c at	g	ctt	gc.	g	ctc	192
															у Ме						
		50						55							0						
aa	a a	tg	gg	c aa	c gg	gC C	:tc	tct	c t	g ga	.c g	ag	gc	c gg	c aa	c c	ett	aco	2	tcc	240
															y As						
6							70						75							80	
ca	a a	a t	gta	ac	c ac	t g	t g	agc	cca	a cc	t ct	c	aaa	ı aa	a ac	c a	ag	tca	1 6	aac	288
															Th						
					8							0						95			
																		•			
a t a	a a	ıc	ctg	gaa	at	a to	c t	gca	ссс	cto	ac	a	gtt	acc	tca	g	aa	gcc	С	t a	336
Ιlε	: As	n .	Leu	Glu	H	e Se	er,	Ala	Pro	Let	ı Th	r '	Val	Thr	Ser	G	l u	Ala	L	eu	
				100				•		105	;					1	10				
act	gt	g g	gc t	gcc	gco	gc	a	cct	cta	atg	gt	C §	gcg	ggc	aac	ac	ca	ctc	a	СС	384
Thr	Va	1 /	lla	Ala	A-l a	Αl	a F	Pro	Leu	Met	Va	l A	la	Gly	Asn	Th	ır	Leu	T	hr	
		1	15						120						125						
atg	ca	a t	са	cag	gcc	cc	g c	ta a	асс	gtg	cac	g	ac	tcc	aaa	c t	t a	agc	a i	t t	432
Met	Gli	n S	er	Gln	Ala	Pro	o L	eu 1	hr	Val	His	Α	รก	Ser	Lve	م آ	11 (lor	ון		

135	140
	135

•	gu	C a	CC	Caa	ı gg	a co	c c	tc	aca	ıgt	g t	ca	gaa	gg	a a	ag	c t a	ı go	C	ctg	caa	480
	Αl	a T	hr	Glr	Gl	y Pi	o L	eu	Thr	· Va	.1 S	er	Glu	Gl	уL	уs	Leu	Al	a]	Leu	Gln	
	14							50						15							160	
	ac	a to	ca	ggc	CC	c ct	c a	cc a	cc	ac	c ga	it a	ıgc	ag	t a	СС	ctt	ac	t a	atc	act	528
						Le																
						16							70							75		
	gco	tc	a	ссс	cct	ct	a ac	t a	c t	gco	ac	t g	gt.	ago	: tt	g	ggc	ati	t g	ac	ttg	576
						Let																0.0
					180						18						-	190			200	
	aaa	ga	g c	сс	a t t	tat	ac	a ca	aa	aat	gg	a a	aa	cta	gg	a c	cta	aag	: t.	ac	ggg	624
						Tyr																024
				95						200							205	_,_		, -	01)	
į	gc t	cct	t	tg	cat	gta	aca	ı ga	.C ;	gac	cta	aa	.c a	act	t t s	z a	CC :	gta	σr	٠a :	ac t	679
						Val																672
		210						21							220		•••	, a i	A1	a	1111	
															220	,						
g	gt	cca	gg	gt g	gtg	act	att	аа	t a	at	act	t c	r t	t or	000	0.4	o f .		4	1 .		500
						Thr																720
	25			- '	-		230	1101	(1	וו ני	1111	JE.			υll	11	ır L	уs	٧a			
							200						Z	35						2	40	

gg	a go	c ti	g g	gt t	tt ga	t tc	a ca	a gg	c aa	t at	g caa	a cti	t aa	t gt	a gca	768
Gl	y Al	a Le	eu G	ly Pl	ne As	p Se	r Gl	n Gl	y As	n Me	t Gli	ı Lei	ı As	n Va	l Ala	
				24	1 5				25	0				25	5	
gg	a gg	a ct	a ag	gg at	t ga	t tc	t ca	a aa	c ag	a cgo	ctt	ata	ct	t ga	t gtt	816
															o Val	
			26					26					270			
														-		
agt	ta	t cc.	gtt	t ga	t gc	t caa	a aac	c caa	a cta	ı aat	cta	aga	c fa	a ppa	cag	864
					p Ala											004
		27					280					285			OIN	
												200				
ggc	cct	c t	tt	t at	a aac	: tca	ı god	сас	ะลลก	tto	o a t	att	220	taa	000	. 010
					e Asn											912
	290					295		. 1115	71511	Dog	300	116	VOII	llyl	ASII	
											500					
aaa	ggc	ctt	tac	: tts	g ttt	ara	ar t	tca	220	a a f	100			- 4 4		000
																960
305	0.,	Dou	1 1 1	bcc	Phe 310	1 11 1	Ala	361	ASII		ser	Lys	Lys	Leu		
000					310					315					320	
at t	222	a t o		4											•	
					gcc			•								1008
vai	ASI	reu	Ser		Ala	Lys	Gly	Leu	Met	Phe	Asp	Ala	Thr	Ala	Ile	
				325					330					335		
gcc																1056
Ala	Ile	Asn	Ala	Gly	Asp	Gly	Leu	Glu	Phe	Gly	Ser :	Pro .	Asn	Ala ·	Pro	

340	345	350
	• • •	

aa	c ac	a aa	at co	c ct	c aa	a ac	a aa	a at	t gg	c ca	t gg	c ct	a ga	a t t	t gat	1104
As	n Th	r As	n Pr	o Le	u Ly	s Th	r Ly	s Ile	e Gl	у Ні	s GI	y Le	u Gl	u Ph	e Asp	
		35	55				36	0				36	5			
tc	a aa	c aa	g gc	t at	g gt	t cc	t aaa	a cta	ı gga	a ac	gg	cct	t ag	t tti	t gac	1152
															e Asp	
	37					379					380				•	
ago	c aca	a gg	t gc	c at	taca	a gta	ı gga	aac	aaa	ı aat	aai	gat	aa⊆	z cta	ı act	1200
															Thr	1200
385					390					395			2,0	, 500	400	
															100	
t t g	tgg	aco	aca	а сса	gct	cca	tct	cct	aac	tgt	ลฐล	cta	aat	σca	gag	1940
				Pro												1248
				405					410	0,0	**** 6	Deu	11311		Glu	
									110					415		
aaa	gat	gc t	ааа	c.t c	act	ffσ	atc	tta	200				4			
																1296
-, -			420	Leu	1111	Leu	Val		IHT	LYS	Cys	Gly		GIn	lle	
			420					425					430			
0 + +	~ · ·		_11													
				tca												1344
Leu	Ala		Val	Ser	Val	Leu	Ala	Val	Lys	Gly	Ser	Leu	Ala	Pro	Ile	
		435					440					445				

į c	t g	ga a	ca	gt	t ca	a ag	gt go	et ca	at c	tt a	tt a	ta ag	ga tt	t ga	ıc g	aa	aat	1392
Se	r G	ly T	'nг	Val	l Gl	n Se	r Al	a Hi	is L	eu II	e I	le Ar	g Ph	e As	p G	lu	Asn	
	45	50					45	55				46	0	•				
gg	a gt	g c	t a	c t a	aa	c aa	t to	c tt	c ct	g ga	.c cc	a ga	a ta	t tg	g a	ac	ttt	1440
													u Ty					
46						47					47						480	
aga	a aa	t gg	ga	gat	ctt	ac	t ga	a gg	c ac	a gc	c ta	t ac	a aac	c gc	t gi	t	gga	1488
													r Ası					
					485					49					49			
ttt	a t į	g cc	t a	aac	cta	tca	gc	t ta	t cc	a aaa	a tc	t cad	c ggt	aaa	a ac	t	gcc	1536
													s Gly					
				500					50					510				
																	•	
aaa	agt	aa	c a	t t	gtc	agt	caa	gtt	tac	: tta	aac	gga	gac	aaa	ac	t	aaa	1584
													Asp					
		51	5					520					525					
cct	gta	aca	a c	t a	acc	att	aca	cta	aac	ggt	aca	cag	gaa	aca	gga	1 8	gac	1632
													Glu					
	530						535					540						
aca	ac t	cca	ag	gtg	gca	tac	tct	atg	tca	t t t	tca	tgg	gac	tgg	tct	g	gc	1680
													Asp					

545					550					555					560	
					gaa Glu											1728
				565					570					575		
	٠															
tac	att	gcc	caa	gaa	ggt	tct	gga	agt	ggt	agt	ggt	t c t	ggc	agc	gga	1776
Tyr	Ile	Ala	Gln	Glu	Gly	Ser	Gly	Ser	Gly	Ser	Gly	Ser	Gly	Ser	Gly	
			580					585					590			
tct	gga	tct	aag	aag	aag	aag	aag	aaa	aag	aag	aaa	aag	aag	aag	aag	1824
Ser	Gly	Ser	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys.	
		595					600					605				
aaa	aaa	aag	aag	aag	aaa	aag	aaa	gga	tcc	tac	tcc	atg	gag	cac	ttc	1872

cgc tgg ggc aag ccg gtg taa 1893
Arg Trp Gly Lys Pro Val
625 630

620

Lys Lys Lys Lys Lys Lys Gly Ser Tyr Ser Met Glu His Phe

615

⟨210⟩ 22

<211> 1923

610

<212> DNA

<213> Artificial Sequence

<220>

<223> DNA coding a modified fiber protein of pWE6.7R-F/asK21MSHb

<220>

<221> CDS

<222> (1).. (1920)

<400> 22

atg aag cgc gca aga ccg tct gaa gat acc ttc aac ccc gtg tat cca 48
Met Lys Arg Ala Arg Pro Ser Glu Asp Thr Phe Asn Pro Val Tyr Pro
1 5 10 15

tat gac acg gaa acc ggt cct cca act gtg cct ttt ctt act cct ccc 96

Tyr Asp Thr Glu Thr Gly Pro Pro Thr Val Pro Phe Leu Thr Pro Pro

20 25 30

ttt gta tcc ccc aat ggg ttt caa gag agt ccc cct ggg gta ctc tct 144
Phe Val Ser Pro Asn Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser

35 40 45

ttg cgc cta tcc gaa cct cta gtt acc tcc aat ggc atg ctt gcg ctc 192 Leu Arg Leu Ser Glu Pro Leu Val Thr Ser Asn Gly Met Leu Ala Leu
50 55 60

aaa	atg	ggo	aac	ggc	ctc	tct	ctg	gac	gag	gcc	ggc	aac	ctt	acc	tcc	240
Lys	Met	Gly	Asn	Gly	Leu	Ser	Leu	Asp	Glu	Ala	Gly	Asn	Leu	Thr	Ser	
65					70		٠			75	•				80	
										,						
caa	aat	gta	acc	act	gtg	agc	cca	cct	ctc	aaa	aaa	acc	aag	tca	aac	288
Gln	Asn	Val	Thr	Thr	Val	Ser	Pro	Pro	Leu	Lys	Lys	Thr	Lys	Ser	Àsn	
				85					90					95		
a t a	aac	ctg	gaa	ata	tct	gca	ccc	ctc	aca	gtt	acc	tca	gaa	gcc	cta	336
Ile	Asn	Leu	Glu	Ile	Ser	Ala	Pro	Leu	Thr	Val	Thr	Ser	Glu	Ala	Leu	
			100					105					110			
						•										
act	gtg	gct	gcc	gcc	gca	c c t	cta	atg	gtc	gcg	ggc	aac	aca	.ctc	acc	384
Thr	Val	Ala	Ala	Ala	Ala	Pro	Leu	Met	Val	Ala	Gly	Asn	Thr	Leu	Thr	
		115					120					125				
atg	caa	tca	cag	gcc	ccg	cta	acc	gtg	cac	gac	tcc	aaa	ctt	agc	a t t	432
Met	Gln	Ser	Gln	Ala	Pro	Leu	Thr	Val	His	Asp	Ser	Lys	Leu	Ser	Ile	
	130					135					140					
gcc	acc	caa	gga	ccc	ctc	aca	gtg	tca	gaa	gga	aag	cta	gcc	ctg	caa	480
Ala	Thr	Gln	Gly	Pro	Leu	Thr	Val	Ser	Glu	Gly	Lys	Leu	Ala	Leu	Gln	
145					150					155					160	
aca	tca	ggc	ссс	ctc	acc	acc	acc	gat	agc	agt	acc	ctt	act	aic	ac t	528
					Thr											

gcc	tca	ccc	cct	c t a	act	ac t	gcc	act	ggt	agc	ttg	ggc	att	gac	ttg	576
Ala	Ser	Pro	Pro	Leu	Thr	Thr	Ala	Thr	Gly	Ser	Leu	Gly	Ile	Asp	Leu	
			180					185					190			
aaa	gag	ccc	att	tat	aca	caa	aat	gga	aaa	cta	gga	cta	aag	tac	ggg	624
Lys	Glu	Pro	Ile	Tyr	Thr	Gln	Asn	Gly	Lys	Leu	Gly	Leu	Lys	Tyr	Gly	
		195					200					205				
gc t	cct	ttg	cat	gta	aca	gac	gac	cta	aac	ac t	ttg	acc	gta	gca	ac t	672
Ala	Pro	Leu	His	Val	Thr	Asp	Asp	Leu	Asn	Thr	Leu	Thr	Val	Ala	Thr	
	210					215					220					
ggt	cca	ggt	gtg	ac t	att	aat	aat	act	tcc	ttg	caa	ac t	aaa	gtt	act	720
Gly	Pro	Gly	Val	Thr	Ile	Asn	Asn	Thr	Ser	Leu	Gln	Thr	Lys	Val	Thr	
225					230					235					240	
gga	gcc	ttg	ggt	t t t	gat	tca	caa	ggc	aat	atg	caa	ctt	aat	gta	gca	768
Gly	Ala	Leu	Gly	Phe	Asp	Ser	Gln	Gly	Asn	Met	Gln	Leu	Asn	Val	Ala	
				245					250					255		
gga	gga	cta	agg	a t t	gat	tct	caa	aac	aga	cgc	ctt	a t a	ctt	gat	gtt	816

Gly Gly Leu Arg Ile Asp Ser Gln Asn Arg Arg Leu Ile Leu Asp Val

•																					
a	gt t	at	cc	g t	tt g	at g	gc t	caa	a aa	ac c	aa	c t	a aa	at (cta	ag	a ci	a g	ga	cag	864
																				Gln	
			27						28							28			•		
																	•				•
gg	gc c	c t	ct	t t t	t a	ta a	ac	tca	gc	СС	ac	aao	c tt	g g	;a t	ati	aa	c t	ac	aac	912
																				Asn	012
		90						295							00		, 110		УI	ASII	
														Ū	00						
aa	a gg	g C	ctt	t a	c t t	g t	tt	aca	gc	t to	a	aac	aa	t t	c c	ааа	ลล	g (*)	t t	നാന	960
					r Le																900
. 30							10						31		C1	Lys	Ly.	S Lt	z u		•
													,01	U						320	
gt	t aa	.C (cta	ago	c ac	t go	c a	ag	ggg	z t t	g :	ato	f f i	t ors	a c	ore t	200	. ~			1000
					Th																1008
					32			•	3	20		330	1110	, ne	ι de	nia	1111			11e	
											Ü	,00						33	อ		
gcc	at	t a	a t	gca	gga	a ga	t g	gg	ctt	σas	a f	f f	aa t	t o		4					
					Gly																1056
				340			ρ 0	l y	LCu			пе	GIY	se	Γŀ	70		ΑI	a I	Pro	
				010						345)						350				
aac	aca		a t	000	a t a																
					ctc																1104
nsn	1 11 1			F I U	Leu	LYS	3 11			He	G I	ly.	His	Gly	y L	eu	Glu	Phe	e A	sp	
		35) 5					3	360						3	65					
,																					
tca																					1152
Ser	Asn	Lу	s A	la	Met	Val	Pr	o L	уs	Leu	Gl	у 7	ſhr	Gly	Le	eu S	Ser	Phe	A:	gp	

370	375	380

ag	c ac	ca į	ggt	gc	c at	t ac	ca gi	a ge	ga aa	ıc aa	a aa	t aa	t ga	t aa	g ct	a act	1200
Se	r Th	ır (Gly	Al	a II	e Th	ır Va	ıl Gl	y As	n Ly	s As	n As	n Ası	Ly:	s Le	u Thr	
38	5					39	0				39	5				400	
t t	g tg	g a	cc	aca	а сс	a gc	t cc	a tc	t cc	t aa	c tg	t ag	a cta	aai	gca	a gag	1248
Le	u Tr	рТ	'h r	Thr	Pr	o Al	a Pr	o Se	r Pr	o Ası	n Cy	s Ar	g Leu	Asn	ı Ala	Glu	
					40	5	_			410)				415	;	
																ata	1296
Lys	S Ası	ΟΑ	la	Lys	Leu	ı Th	r Lei	ı Va	l Lei	ı Thr	Lys	Cys	Gly	Ser	Gln	Ile	
				420					425	5				430			
ctt	gc t	a (ca	gtt	tca	gtt	. tte	gct	gtt	aaa	ggc	agt	ttg	gc t	cca	ata	1344
			_														
Leu	Ala			Val	Ser	Val	Leu	Ala	. Val	Lys	Gly	Ser	Leu	Ala	Pro	Ile	
		43	35					440					445				
4 - 4																	
													ttt				1392
ser		Th	г \	al	Gln	Ser	Ala	His	Leu	Ile	Ile	Arg	Phe	Asp	Glu	Asn	
	450						455					460					
	,	i															
													tat				1440
	vai	Lei	1 L	eu ,			Ser	Phe	Leu	Asp	Pro	Glu	Tyr '	Trp.	Asn	Phe	
465						470					475					480	

aga	aat	gg	a ga	t ct	t ac	t ga	a gg	c ac	a gc	c ta	t ac	a aac	gc	t gt	t ggs	a 1488
												r Asr				
				48					490					49		,
														10	U	
ttt	atg	cct	aac	c t	a tc	a gc	t ta	t cca	ı aaa	ı tc	t ca	c ggt	aaa	a ac	t gcc	1536
												s Gly				
			500					505					510			•
•													010	,		
aaa	agt	aac	att	gto	c ag	t caa	a gtt	tac	t t a	aac	gga	ı gac	aaa	acı	t aaa	1584
												Asp				
		515					520					525			_,_	
cct	gta	aca	cta	acc	att	aca	cta	aac	ggt	aca	cag	gaa	aca	gga	gac	1632
Pro V																
	30					535					540				,	
aca a	.ct (cca	agt	gca	tac	t c t	atg	tca	ttt	tca	tgg	gac	t gg	tct	ggc	1680
Thr T																1000
545					550					555		•		501	560	
															000	
cac a	ac t	ac	att	aat	gaa	ata	ttt	gcc	acc	tcg	agt	tac	act	t t t	tca	1790
His A																1728
				565					570	~~.	501	7 7 1			361	
				•.				,						575		
tac at	t g	cc c	aa g	gaa	cca	tca	gcc	tcc g	gca	tct	gc t	tcc g	gcc (cct	gga	1776

Tyr Ile Ala Gln Glu Pro Ser Ala Ser Ala Ser Ala Ser Ala Pro Gly
580 585 590

aaa aag aag aag aag aaa aag gaa tcc gcc gag aag aag gac gag 1872 Lys Lys Lys Lys Lys Lys Lys Gly Ser Ala Glu Lys Lys Asp Glu 610 615 620

ggc ccc tac agg atg gag cac ttc cgc tgg ggc agc ccg ccc aag gac 1920 Gly Pro Tyr Arg Met Glu His Phe Arg Trp Gly Ser Pro Pro Lys Asp 625 630 635 640

<210> 23

<211> 1923

<212> DNA

<213> Artificial Sequence

<220>

<223> DNA coding a modified fiber protein of pWE6. 7R-F/gsK21MSHb

<220>

<221> CDS <222> (1).. (1920)

<400> 23

atg	aag	cgc	gca	aga	ccg	tct	gaa	gat	acc	t t c	aac	ссс	gtg	tat	cca	48
Met	Lys	Arg	Ala	Arg	Pro	Ser	Glu	Asp	Thr	Phe	Asn	Pro	Val	Tyr	Pro	
1				5					10					15		

tat gac acg gaa acc ggt cct cca act gtg cci ttt ctt act cct ccc 96

Tyr Asp Thr Glu Thr Gly Pro Pro Thr Val Pro Phe Leu Thr Pro Pro

20 25 30

ttt gta tcc ccc aat ggg ttt caa gag agt ccc cct ggg gta ctc tct 144
Phe Val Ser Pro Asn Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
35 40 45

ttg cgc cta tcc gaa cct cta gtt acc tcc aat ggc atg ctt gcg ctc 192 Leu Arg Leu Ser Glu Pro Leu Val Thr Ser Asn Gly Met Leu Ala Leu 50 55 60

aaa atg ggc aac ggc ctc tct ctg gac gag gcc ggc aac ctt acc tcc 240 Lys Met Gly Asn Gly Leu Ser Leu Asp Glu Ala Gly Asn Leu Thr Ser 65

caa aat gta acc act gtg agc cca cct ctc aaa aaa acc aag tca aac 288 Gln Asn Val Thr Thr Val Ser Pro Pro Leu Lys Lys Thr Lys Ser Asn

. 85	90	95

αι	a ad	וט טו	g ga	ia ai	a tc	t gc	a cc	c ct	c ac	a gt	t ac	c tc	a ga	a gc	c cta	336
I 1	e As	n Le	eu Gl	u II	e Se	r Al	a Pr	o Le	u Th	r Va	l Th	r Se	r Gl	u Ala	a Leu	
			10					10					11			
ac	t gt	g gc	t gc	c gc	c gc	а сс	t ct	a at s	ggt	c gc	g gg(c aac	aca	a cto	acc	384
															ı Thr	
		11					120					125				
•																
a t g	g caa	a to	a ca	g gc	cce	g cta	aco	gtg	cac	gao	c tec	aaa	. cti	t ago	att	432
											Ser					
	130					135					140					
															-	
gcc	acc	caa	gga	ссс	ctc	aca	gtg	tca	gaa	gga	aag	cta	gcc	ctg	caa	480
											Lys					
145					150					155					160	
aca	tca	ggc	ссс	ctc	acc	acc	acc	gat	agc	agt	acc	ctt	act	atc	act	528
											Thr					
				165					170					175		
gcc	tca	ccc	cct	cta	act	act	gcc	ac t	ggt	agc	ttg	ggc	att	gac	ttg	576
											Leu					
			180					185					190			

aa	a ga	ıg c	cc a	att	t a	t ac	a c	aa	aa	t gg	a aa	aa c	t a	gga	cta	a a a	g t	ac	ggg	624
Ly	s Gl	u P	ro]	lle	Ту	r Th	ır G	ln	Ası	n Gl	y Ly	's L	eu	Gly	Lei	ı Ly	s T	уг	Gly-	
			95						200						205					
																			,	
gct	t cc	t t	tg c	at	gta	a ac	a g	ac	gao	ct	a aa	.c a	c t	ttg	acc	gt	a g	сa	act	672
											u As									
	21							15						220						
											•						•			
gg t	cca	a gg	ţt g	tg	ac t	at	t a	a t	aat	ac	t tc	c tt	g	caa	act	aaa	a gi	t t	act	720
											r Se									, 20
225						23						23				•		-	240	
										٠									210	
gga	gcc	: tt	g g	gt	ttt	ga	t to	a	caa	ggo	aa	t at	g c	aa	ctt	aat	gt	a	gca	768
Gly																				
					245						250						25		u	
																	- 0	Ĭ		
gga	gga	cta	a ag	gg a	att	gat	t c	t c	caa	aac	aga	cg	СС	t t	ata	ctt	ga	t :	gtt	816
Gly																				
			26							265						270		F		
																•				
agt	tat	cce	; tt	t g	at	gct	caa	a a	.ac	caa	cta	aat	c i	ta a	aga	cta	øσ	a d	າສອ	864
Ser																				004
		275							80						285	LCu	01	y	3111	
								_	- •					2	.00					
ggc (cct	ctt	t t :	t a	t a	aac	tea	σı	C C	cac	220	tta	αn	ı f	. + +	0.0.0				0.1.0
Sly I																				912
			- 41 (1111		_			~ > 11	ا و دسو ر	40.0	11 1	10		1 7	. ^	~	

290	295	300

aa	a gg	c ci	it ta	c tt	gtt	t ac	a gc	t tc	a aa	c aa	t tc	c aa	a aa	gct	t gag	960
															u Glu	
30					31					31					320	
gt	t aa	c ct	a ag	c ac	t gc	c aa	g ggg	gttg	g atg	g tti	t ga	c gci	aca	ı gc	e ata	1008
															ı Ile	1000
				32					330					335		
gcc	at	t aa	t gca	a gga	a ga	t ggg	ctt	gaa	ı ttt	ggt	tica	a cct	aat	gca	cca	1056
						o Gly										1000
			340					345					350	1110	. 110	
aac	aca	aat	ccc	cto	aaa	ı aca	aaa	att	ggc	cat	ggc	cta	gaa	† † †	σat	1104
						Thr										1104
		355					360				01,	365	oru	1110	пор	
												000				
tca	aac	aag	gct	atg	gtt	cct	aaa	.c.ta	gga	act	aac	c t t	n cr f	+ + +	~ ~ ~	1150
						Pro										1152
	370					375	2,5	DCu	dly	1 11 1		Leu	ser	Рпе	ASP	
						0.0					380					
agc	аса	ggt	σcc	att	202	a t a	~~~									
						gta										1200
885	- 111	o ry	111 CL	116		v d I	σιу	ASN			Asn	Asp	Lys	Leu	Thr	
, , ,					390			•		395					400	

t t	g te	g a	cc a	.ca	cca	gc	t cc	a tc	t cc	t aa	c tg	t ag	a cta	a aa	t gc	a gag	1248
Le	u Tr	рТ	hr T	hr	Pro	Al	a Pro	o Se	r Pr	o As	n Cy	s Ar	g Lei	ı As	n Al	a Glu	
					405					41	0			•	41	5	
aa	a ga	t go	ct a	aa	ctc	ac	tttg	ggt	c t t	a aca	a aa	a tg	t ggo	ag	t caa	a ata	1296
Ly	s As	p A	la L	уS	Leu	Thi	r Lei	ı Va	l Le	u Th	r Ly	s Cy:	s Gly	, Se	r Gli	ı Ile	,
			4	20					42	5				430)		
c t	t gc	t ac	a g	t t	tca	gtt	ttg	gci	t gt	t aaa	ı ggo	agi	ttg	gct	cca	ata	1344
Lei	ı Ala	a Th	r Va	al :	Ser	Val	Leu	Ala	a Val	Lys	Gly	/ Sei	Leu	Ala	n Pro	Ile	
		43	5					440)				445				
tct	gga	ac	a gt	t	caa	agt	gc t	cat	ctt	att	ata	aga	ttt	gac	gaa	aat	1392
Ser	Gly	Th	r Va	1 (Gln	Ser	Ala	His	Leu	Ile	Ile	Arg	Phe	Asp	Glu	Asn	
	450	}					455					460					
gga	gtg	ct	a ct	a a	ac	aat	tcc	t t c	ctg	gac	cca	gaa	tat	t gg	aac	ttt	1440
Gly	Val	Lei	ı Le	u A	sn	Asn	Ser	Phe	Leu	Asp	Pro	Glu	Tyr	Trp	Asn	Phe	
465						470			•	•	475					480	
aga	aat	gga	ga	t c	t t	ac t	gaa	ggc	aca	gcc	tat	aca	aac	gc t	gtt	gga	1488
Arg	Asn	Gly	' Ası	o L	eu ′	Thr	Glu	Gly	Thr	Ala	Tyr	Thr	Asn	Ala	Val	Gly	
				48	85					490					495		
														,			
t t t	atg	cct	aac	ci	ta t	ca	gc t	tat	cca	aaa	tct	cac	ggt	aaa	act	gcc	1536
													Gly				

500	505	510

aaa	agt	aac	ati	gto	agt	caa	gtt	tac	tta	aac	gga	gac	aaa	act	aaa	1584
Lys	Ser	Asn	Ile	. Val	Ser	Gln	Val	Tyr	Leu	Asn	Gly	Asp	Lys	Thr	Lys	
		515	;				520	ı				525				
cct	gta	aca	cta	acc	att	aca	cta	aac	ggt	aca	cag	gaa	aca	gga	gac	1632
Pro	Val	Thr	Leu	Thr	Ile	Thr	Leu	Asn	Gly	Thr	Gln	Glu	Thr	Gly	Asp	
	530					535					540			٠		
aca	act	cca	agt	gca	tac	tct	atg	tca	ttt	tca	· tgg	gac	t gg	tct	ggc	1680
Thr	Thr	Pro	Ser	Ala	Tyr	Ser	Met	Ser	Phe	Ser	Trp	Asp	Trp	Ser	Gly	
545	-				550					555					560	
cac	aac	tac	att	aat	gaa	ata	ttt	gcc	acc	tcg	agt	tac	act	ttt	tca	1728
His	Asn	Tyr	Ile	Asn	Glu	Ile	Phe	Ala	Thr	Ser	Ser	Tyr	Thr	Phe	Ser	
				565					570					575		
tac	a t t	gcc	caa	gaa	ggt	tct	gga	agt	ggt	agt	ggt	tct	ggc	agc	gga	1776
						Ser										
			580					585					590		•	
tct	gga	tct	aag	aag	aag	aag	aag	aaa	aag	aag	aaa	aag	aag	aag	aag	1824
						Lys										
		595					600	-	-			605	-,0	-,0	~ , 0	

aaa aaa aag aag aag aaa aag aaa gga tcc gcc gag aag aag gac gag 1872 Lys Lys Lys Lys Lys Lys Lys Gly Ser Ala Glu Lys Lys Asp Glu 610 615 620

ggc ccc tac agg atg gag cac ttc cgc tgg ggc agc ccg ccc aag gac 1920 Gly Pro Tyr Arg Met Glu His Phe Arg Trp Gly Ser Pro Pro Lys Asp 625 630 635 640

taa 1923

<210> 24

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> DNA coding AS linker

<220>

<221> CDS

⟨222⟩ (1).. (33)

<400> 24

cca tca gcc tcc gca tct gct tcc gcc cct gga
Pro Ser Ala Ser Ala Ser Ala Ser Ala Pro Gly
1
5
10

33

<210> 25

<211> 11

<212> PRT

<213> Artificial Sequence

<220>

<223> AS linker peptide

<400> 25

Pro Ser Ala Ser Ala Ser Ala Pro Gly

1

5

10

<210> 26

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> DNA coding GS linker

<220>

<221> CDS

<222> (1).. (33)

```
<400> 26
```

ggt tct gga agt ggt agt ggt tct ggc agc gga Gly Ser Gly Ser Gly Ser Gly Ser Gly 1 5 10

33

<210> 27

<211> 11

<212> PRT

<213> Artificial Sequence

<220>

<223> GS linker peptide

<400> 27 ·

Gly Ser Gly Ser Gly Ser Gly Ser Gly

1 5 10

<210> 28

<211> 108

<212> DNA

<213> Artificial Sequence

<220>

<223> DNA coding asK21 linker

<220>

<221> CDS

<222> (1).. (108)

<400> 28

cca tca gcc tcc gca tct gct tcc gcc cct gga tct gga tct aag aag

Pro Ser Ala Ser Ala Ser Ala Ser Ala Pro Gly Ser Gly Ser Lys Lys

1 5 10 15

aaa aag aaa gga Lys Lys Lys Gly

35

<210> 29

<211> 36

<212> PRT

<213> Artificial Sequence

<220>

<223> asK21 linker peptide

<400> 29

Pro Ser Ala Ser Ala Ser Ala Ser Ala Pro Gly Ser Gly Ser Lys Lys

1 10 15

Lys Lys Lys Gly

35

<210> 30

<211> 108

<212> DNA

<213> Artificial Sequence

<220>

<223> DNA coding gsK21 linker

<220>

<221> CDS

<222> (1).. (108)

<400> 30

1

ggt tct gga agt ggt agt ggt tct ggc agc gga tct gga tct aag aag 48 Gly Ser Gly Ser Gly Ser Gly Ser Gly Ser Gly Ser Lys Lys 1 5 10 15 96 20 25 30 aaa aag aaa gga 🙃 108 Lys Lys Lys Gly .35 <210> 31 <211> 36 <212> PRT <213> Artificial Sequence <220> <223> gsK21 linker peptide <400> 31 Gly Ser Gly Ser Gly Ser Gly Ser Gly Ser Gly Ser Lys

5

10

15

LI

į.

Lys Lys Lys Gly

35

<210> 32

<211> 605

<212> PRT

<213> Artificial Sequence

<220>

<223> a modified fiber protein encoded in pWE6.7R-F/asMSHa

<400> 32

Met Lys Arg Ala Arg Pro Ser Glu Asp Thr Phe Asn Pro Val Tyr Pro 1 5 10 15

Tyr Asp Thr Glu Thr Gly Pro Pro Thr Val Pro Phe Leu Thr Pro Pro
20 25 30

Phe Val Ser Pro Asn Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser

Leu Arg Leu Ser Glu Pro Leu Val Thr Ser Asn Gly Met Leu Ala Leu
50 55 60

Lys	Me t	Gly	Asn	Gly	Leu	Ser	Leu	Asp	Glu	Ala	Gly	Asn	Leu	Thr	Ser
65					70					75		·			80

Gln Asn Val Thr Thr Val Ser Pro Pro Leu Lys Lys Thr Lys Ser Asn 85 90 95

Ile Asn Leu Glu Ile Ser Ala Pro Leu Thr Val Thr Ser Glu Ala Leu 100 105 110

Thr Val Ala Ala Ala Pro Leu Met Val Ala Gly Asn Thr Leu Thr
115 120 125

Met Gln Ser Gln Ala Pro Leu Thr Val His Asp Ser Lys Leu Ser Ile 130 135 140

Ala Thr Gln Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu Gln 145 150 155 160

Thr Ser Gly Pro Leu Thr Thr Thr Asp Ser Ser Thr Leu Thr Ile Thr
165 170 175

Ala Ser Pro Pro Leu Thr Thr Ala Thr Gly Ser Leu Gly Ile Asp Leu 180 185 190

Lys Glu Pro Ile Tyr Thr Gln Asn Gly Lys Leu Gly Leu Lys Tyr Gly
195 200 205

Ala Pro Leu 1	1115	vai	ınr	Asp	Asp	Leu	Asn	Thr	Leu	Thr	Val	Ala	Thr
210	٠			215					220				

Gly Pro Gly Val Thr Ile Asn Asn Thr Ser Leu Gln Thr Lys Val Thr 225 230 235 240

Gly Ala Leu Gly Phe Asp Ser Gln Gly Asn Met Gln Leu Asn Val Ala
245
250
255

Gly Gly Leu Arg Ile Asp Ser Gln Asn Arg Arg Leu Ile Leu Asp Val 260 265 270

Ser Tyr Pro Phe Asp Ala Gin Asn Gin Leu Asn Leu Arg Leu Gly Gin 275 280 285

Gly Pro Leu Phe Ile Asn Ser Ala His Asn Leu Asp Ile Asn Tyr Asn 290 295 300

Lys Gly Leu Tyr Leu Phe Thr Ala Ser Asn Asn Ser Lys Lys Leu Glu 305 310 315 320

Val Asn Leu Ser Thr Ala Lys Gly Leu Met Phe Asp Ala Thr Ala Ile 325 330 335

Ala Ile Asn Ala Gly Asp Gly Leu Glu Phe Gly Ser Pro Asn Ala Pro

Asn Thr Asn Pro Leu Lys Thr Lys Ile Gly His Gly Leu Glu Phe Asp 355 360 365

Ser Asn Lys Ala Met Val Pro Lys Leu Gly Thr Gly Leu Ser Phe Asp 370 375 380

Ser Thr Gly Ala Ile Thr Val Gly Asn Lys Asn Asn Asp Lys Leu Thr 385 390 395 400

Leu Trp Thr Thr Pro Ala Pro Ser Pro Asn Cys Arg Leu Asn Ala Glu
405 410 415

Lys Asp Ala Lys Leu Thr Leu Val Leu Thr Lys Cys Gly Ser Gln Ile
420
430

Leu Ala Thr Val Ser Val Leu Ala Val Lys Gly Ser Leu Ala Pro Ile 435 440 445

Ser Gly Thr Val Gln Ser Ala His Leu Ile Ile Arg Phe Asp Glu Asn 450 455 460

Gly Val Leu Leu Asn Asn Ser Phe Leu Asp Pro Glu Tyr Trp Asn Phe 465 470 475 480 Arg Asn Gly Asp Leu Thr Glu Gly Thr Ala Tyr Thr Asn Ala Val Gly
485
490
495

Phe Met Pro Asn Leu Ser Ala Tyr Pro Lys Ser His Gly Lys Thr Ala
500 505 510

Lys Ser Asn Ile Val Ser Gin Val Tyr Leu Asn Gly Asp Lys Thr Lys
515 520 525

Pro Val Thr Leu Thr Ile Thr Leu Asn Gly Thr Gln Glu Thr Gly Asp 530 535 540

Thr Thr Pro Ser Ala Tyr Ser Met Ser Phe Ser Trp Asp Trp Ser Gly 545 550 550 560

His Asn Tyr Ile Asn Glu Ile Phe Ala Thr Ser Ser Tyr Thr Phe Ser 565 570 575

Tyr Ile Ala Gln Glu Pro Ser Ala Ser Ala Ser Ala Ser Ala Pro Gly
580 585 590

Ser Tyr Ser Met Glu His Phe Arg Trp Gly Lys Pro Val
595 600 605

<210> 33

<211> 615 <212> PRT <213> Artificial Sequence <220> <223> a modified fiber protein encoded in pWE6.7R-F/asMSHb <400> 33 Met Lys Arg Ala Arg Pro Ser Glu Asp Thr Phe Asn Pro Val Tyr Pro 1 5 10 15 Tyr Asp Thr Glu Thr Gly Pro Pro Thr Val Pro Phe Leu Thr Pro Pro 20 25 30 Phe Val Ser Pro Asn Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser 35 40 45 Leu Arg Leu Ser Glu Pro Leu Val Thr Ser Asn Gly Met Leu Ala Leu 50 55 60 Lys Met Gly Asn Gly Leu Ser Leu Asp Glu Ala Gly Asn Leu Thr Ser 65 70 75 80

90 `

95

Gln Asn Val Thr Thr Val Ser Pro Pro Leu Lys Lys Thr Lys Ser Asn

85

Ile	Asn	Leu	Glu	Ile	Ser	Ala	Pro	Leu	Thr	Val	Thr	Ser	Glu	Ala	Leu
			100					105					110		

Thr Val Ala Ala Ala Pro Leu Met Val Ala Gly Asn Thr Leu Thr 115 120 125

Met Gln Ser Gln Ala Pro Leu Thr Val His Asp Ser Lys Leu Ser Ile 130 135 140

Ala Thr Gln Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu Gln 145 150 155 160

Thr Ser Gly Pro Leu Thr Thr Thr Asp Ser Ser Thr Leu Thr Ile Thr 165 170 175

Ala Ser Pro Pro Leu Thr Thr Ala Thr Gly Ser Leu Gly Ile Asp Leu 180 185 190

Lys Glu Pro Ile Tyr Thr Gln Asn Gly Lys Leu Gly Leu Lys Tyr Gly
195 200 205

Ala Pro Leu His Val Thr Asp Asp Leu Asn Thr Leu Thr Val Ala Thr 210 215 220

Gly Pro Gly Val Thr Ile Asn Asn Thr Ser Leu Gln Thr Lys Val Thr 225 230 235 240

Gly	Ala	Leu	Gly	Phe	Asp	Ser	Gln	Gly	Asn	Met	Gln	Leu	Asn	Val	Ala
				245					250					255	

Gly Gly Leu Arg Ile Asp Ser Gln Asn Arg Arg Leu Ile Leu Asp Val 260 265 270

Ser Tyr Pro Phe Asp Ala Gln Asn Gln Leu Asn Leu Arg Leu Gly Gln
275
280
285

Gly Pro Leu Phe Ile Asn Ser Ala His Asn Leu Asp Ile Asn Tyr Asn 290 295 300

Lys Gly Leu Tyr Leu Phe Thr Ala Ser Asn Asn Ser Lys Lys Leu Glu 305 310 315 320

Val Asn Leu Ser Thr Ala Lys Gly Leu Met Phe Asp Ala Thr Ala Ile 325 330 335

Ala Ile Asn Ala Gly Asp Gly Leu Glu Phe Gly Ser Pro Asn Ala Pro 340 345 350

Asn Thr Asn Pro Leu Lys Thr Lys Ile Gly His Gly Leu Glu Phe Asp 355 360 365

Ser Asn Lys Ala Met Val Pro Lys Leu Gly Thr Gly Leu Ser Phe Asp

CJ

The state of the s

LII

ļ.,

W. H. H.

370 -	375	380

Ser Thr Gly Ala Ile Thr Val Gly Asn Lys Asn Asn Asp Lys Leu Thr 385 390 395 400

Leu Trp Thr Thr Pro Ala Pro Ser Pro Asn Cys Arg Leu Asn Ala Glu
405 410 415

Lys Asp Ala Lys Leu Thr Leu Val Leu Thr Lys Cys Gly Ser Gln Ile
420
430

Leu Ala Thr Val Ser Val Leu Ala Val Lys Gly Ser Leu Ala Pro Ile 435 440 445

Ser Gly Thr Val Gln Ser Ala His Leu Ile Ile Arg Phe Asp Glu Asn 450 455 460

Gly Val Leu Leu Asn Asn Ser Phe Leu Asp Pro Glu Tyr Trp Asn Phe
465 470 475 480

Arg Asn Gly Asp Leu Thr Glu Gly Thr Ala Tyr Thr Asn Ala Val Gly
485 490 495

Phe Met Pro Asn Leu Ser Ala Tyr Pro Lys Ser His Gly Lys Thr Ala 500 505 510 Lys Ser Asn Ile Val Ser Gln Val Tyr Leu Asn Gly Asp Lys Thr Lys
515 520 525

Pro Val Thr Leu Thr Ile Thr Leu Asn Gly Thr Gln Glu Thr Gly Asp 530 535 540

Thr Thr Pro Ser Ala Tyr Ser Met Ser Phe Ser Trp Asp Trp Ser Gly
545 550 555 560

His Asn Tyr Ile Asn Glu Ile Phe Ala Thr Ser Ser Tyr Thr Phe Ser 565 570 575

Tyr Ile Ala Gln Glu Pro Ser Ala Ser Ala Ser Ala Ser Ala Pro Gly
580 585 590

Ser Ala Glu Lys Lys Asp Glu Gly Pro Tyr Arg Met Glu His Phe Arg 595 600 605

Trp Gly Ser Pro Pro Lys Asp 610 615

<210> 34

<211> 605

<212> PRT

<213> Artificial Sequence

<220>

<223> a modified fiber protein encoded in pWE6.7R-F/gsMSHa

<400> 34

Met Lys Arg Ala Arg Pro Ser Glu Asp Thr Phe Asn Pro Val Tyr Pro 1 5 10 15

Tyr Asp Thr Glu Thr Gly Pro Pro Thr Val Pro Phe Leu Thr Pro Pro 20 25 30

Phe Val Ser Pro Asn Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser 35 40 45

Leu Arg Leu Ser Glu Pro Leu Val Thr Ser Asn Gly Met Leu Ala Leu 50 55 60

Lys Met Gly Asn Gly Leu Ser Leu Asp Glu Ala Gly Asn Leu Thr Ser 65 70 75 80

Gln Asn Val Thr Thr Val Ser Pro Pro Leu Lys Lys Thr Lys Ser Asn 85 . 90 95

Ile Asn Leu Glu Ile Ser Ala Pro Leu Thr Val Thr Ser Glu Ala Leu 100 105 110 Thr Val Ala Ala Ala Pro Leu Met Val Ala Gly Asn Thr Leu Thr
115 120 125

Met Gln Ser Gln Ala Pro Leu Thr Val His Asp Ser Lys Leu Ser Ile 130 135 140

Ala Thr Gln Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu Gln 145 150 155 160

Thr Ser Gly Pro Leu Thr Thr Asp Ser Ser Thr Leu Thr Ile Thr
165 170 175

Ala Ser Pro Pro Leu Thr Thr Ala Thr Gly Ser Leu Gly Ile Asp Leu 180 185 190

Lys Glu Pro Ile Tyr Thr Gln Asn Gly Lys Leu Gly Leu Lys Tyr Gly
195 200 205

Ala Pro Leu His Val Thr Asp Asp Leu Asn Thr Leu Thr Val Ala Thr 210 215 220

Gly Pro Gly Val Thr Ile Asn Asn Thr Ser Leu Gln Thr Lys Val Thr 225 230 235 240

Gly Ala Leu Gly Phe Asp Ser Gln Gly Asn Met Gln Leu Asn Val Ala
245
250
255

Gly Gly Leu Arg Ile Asp Ser Gln Asn Arg Arg Leu Ile Leu Asp Val 260 265 270

Ser Tyr Pro Phe Asp Ala Gln Asn Gln Leu Asn Leu Arg Leu Gly Gln
275
280
285

Gly Pro Leu Phe Ile Asn Ser Ala His Asn Leu Asp Ile Asn Tyr Asn 290 295 300

Lys Gly Leu Tyr Leu Phe Thr Ala Ser Asn Asn Ser Lys Lys Leu Glu 305 310 315 320

Val Asn Leu Ser Thr Ala Lys Gly Leu Met Phe Asp Ala Thr Ala Ile 325 330 335

Ala Ile Asn Ala Gly Asp Gly Leu Glu Phe Gly Ser Pro Asn Ala Pro 340 345 350

Asn Thr Asn Pro Leu Lys Thr Lys Ile Gly His Gly Leu Glu Phe Asp 355 360 365

Ser Asn Lys Ala Met Val Pro Lys Leu Gly Thr Gly Leu Ser Phe Asp 370 375 380

Ser Thr Gly Ala Ile Thr Val Gly Asn Lys Asn Asn Asp Lys Leu Thr

þ.i.

385	390	395	400

Leu Trp Thr Thr Pro Ala Pro Ser Pro Asn Cys Arg Leu Asn Ala Glu
405 410 415

Lys Asp Ala Lys Leu Thr Leu Val Leu Thr Lys Cys Gly Ser Gln Ile
420 425 430

Leu Ala Thr Val Ser Val Leu Ala Val Lys Gly Ser Leu Ala Pro Ile 435 440 445

Ser Gly Thr Val Gln Ser Ala His Leu Ile Ile Arg Phe Asp Glu Asn 450 455 460

Gly Val Leu Leu Asn Asn Ser Phe Leu Asp Pro Glu Tyr Trp Asn Phe 465 470 475 480

Arg Asn Gly Asp Leu Thr Glu Gly Thr Ala Tyr Thr Asn Ala Val Gly
485
490
495

Phe Met Pro Asn Leu Ser Ala Tyr Pro Lys Ser His Gly Lys Thr Ala
500 505 510

Lys Ser Asn Ile Val Ser Gln Val Tyr Leu Asn Gly Asp Lys Thr Lys
515 520 525

Pro Val Thr Leu Thr Ile Thr Leu Asn Gly Thr Gln Glu Thr Gly Asp
530 535 540

Thr Thr Pro Ser Ala Tyr Ser Met Ser Phe Ser Trp Asp Trp Ser Gly 545 550 550 560

His Asn Tyr Ile Asn Glu Ile Phe Ala Thr Ser Ser Tyr Thr Phe Ser 565 570 575

Tyr Ile Ala Gln Glu Gly Ser Gly Ser Gly Ser Gly Ser Gly Ser Gly Ser Gly 580 590

Ser Tyr Ser Met Glu His Phe Arg Trp Gly Lys Pro Val
595 600 605

<210> 35

<211> 615

<212> PRT

<213> Artificial Sequence

<220>

<223> a modified fiber protein encoded in pWE6.7R-F/gsMSHb

<400> 35

Met Lys Arg Ala Arg Pro Ser Glu Asp Thr Phe Asn Pro Val Tyr Pro

C.)

þ.5

design in the last that the

Tyr	Asp	Thr	Glu	Thr	Gly	Pro	Pro	Thr	Val	Pro	Phe	Leu	Thr	Pro	Pro
			20					25					30		

Phe Val Ser Pro Asn Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser 35 40 45

Leu Arg Leu Ser Glu Pro Leu Val Thr Ser Asn Gly Met Leu Ala Leu 50 55 60

Lys Met Gly Asn Gly Leu Ser Leu Asp Glu Ala Gly Asn Leu Thr Ser 65 70 75 80

Gln Asn Val Thr Thr Val Ser Pro Pro Leu Lys Lys Thr Lys Ser Asn 85 90 95

Ile Asn Leu Glu Ile Ser Ala Pro Leu Thr Val Thr Ser Glu Ala Leu 100 105 110

Thr Val Ala Ala Ala Ala Pro Leu Met Val Ala Gly Asn Thr Leu Thr 115 120 125

Met Gln Ser Gln Ala Pro Leu Thr Val His Asp Ser Lys Leu Ser Ile 130 135 140

Ala	Thr	Gln	Gly	Pro	Leu	Thr	Val	Ser	Glu	Gly	Lys	Leu	Ala	Leu	Gln
145					150					155					160

Thr Ser Gly Pro Leu Thr Thr Asp Ser Ser Thr Leu Thr Ile Thr
165 170 175

Ala Ser Pro Pro Leu Thr Thr Ala Thr Gly Ser Leu Gly Ile Asp Leu 180 185 190

Lys Glu Pro Ile Tyr Thr Gln Asn Gly Lys Leu Gly Leu Lys Tyr Gly
195 200 205

Ala Pro Leu His Val Thr Asp Asp Leu Asn Thr Leu Thr Val Ala Thr 210 215 220

Gly Pro Gly Val Thr Ile Asn Asn Thr Ser Leu Gln Thr Lys Val Thr 225 230 235 240

Gly Ala Leu Gly Phe Asp Ser Gln Gly Asn Met Gln Leu Asn Val Ala
245
250
255

Gly Gly Leu Arg Ile Asp Ser Gln Asn Arg Arg Leu Ile Leu Asp Val 260 265 270

Ser Tyr Pro Phe Asp Ala Gln Asn Gln Leu Asn Leu Arg Leu Gly Gln 275 280 285

Gly Pro Leu Phe Ile Asn Ser Ala His Asn Leu Asp Ile Asn Tyr Asn 290 295 300

Lys Gly Leu Tyr Leu Phe Thr Ala Ser Asn Asn Ser Lys Lys Leu Glu 305 310 315 320

Val Asn Leu Ser Thr Ala Lys Gly Leu Met Phe Asp Ala Thr Ala Ile 325 330 335

Ala Ile Asn Ala Gly Asp Gly Leu Glu Phe Gly Ser Pro Asn Ala Pro 340 345 350

Asn Thr Asn Pro Leu Lys Thr Lys Ile Gly His Gly Leu Glu Phe Asp 355 360 365

Ser Asn Lys Ala Met Val Pro Lys Leu Gly Thr Gly Leu Ser Phe Asp 370 375 380

Ser Thr Gly Ala Ile Thr Val Gly Asn Lys Asn Asn Asp Lys Leu Thr 385 390 395 400

Leu Trp Thr Thr Pro Ala Pro Ser Pro Asn Cys Arg Leu Asn Ala Glu
405 410 415

Lys Asp Ala Lys Leu Thr Leu Val Leu Thr Lys Cys Gly Ser Gln Ile

Leu Ala Thr Val Ser Val Leu Ala Val Lys Gly Ser Leu Ala Pro Ile 435 440 445

Ser Gly Thr Val Gln Ser Ala His Leu Ile Ile Arg Phe Asp Glu Asn 450 455 460

Gly Val Leu Leu Asn Asn Ser Phe Leu Asp Pro Glu Tyr Trp Asn Phe 465 470 475 480

Arg Asn Gly Asp Leu Thr Glu Gly Thr Ala Tyr Thr Asn Ala Val Gly
485
490
495

Phe Met Pro Asn Leu Ser Ala Tyr Pro Lys Ser His Gly Lys Thr Ala
500 505 510

Lys Ser Asn Ile Val Ser Gln Val Tyr Leu Asn Gly Asp Lys Thr Lys
515 520 525

Pro Val Thr Leu Thr Ile Thr Leu Asn Gly Thr Gln Glu Thr Gly Asp 530 535 540

Thr Thr Pro Ser Ala Tyr Ser Met Ser Phe Ser Trp Asp Trp Ser Gly
545 550 555 560

His Asn Tyr Ile Asn Glu Ile Phe Ala Thr Ser Ser Tyr Thr Phe Ser
565 570 575

Tyr Ile Ala Gln Glu Gly Ser Gl

Ser Ala Glu Lys Lys Asp Glu Gly Pro Tyr Arg Met Glu His Phe Arg 595 600 605

Trp Gly Ser Pro Pro Lys Asp 610 615

<210> 36

<211> 630

<212> PRT

<213> Artificial Sequence

<220>

<223> a modified fiber protein encoded in pWE6.7R-F/asK21MSHa

<400> 36

Met Lys Arg Ala Arg Pro Ser Glu Asp Thr Phe Asn Pro Val Tyr Pro

1 5 10 15

Tyr Asp Thr Glu Thr Gly Pro Pro Thr Val Pro Phe Leu Thr Pro Pro

[]
f ,[#
V.)
ļ.i.
450
₽.b
LII
ļ. i
韗
C)
22.52
C)
ar i

Phe Val Ser Pro Asn Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser

35 40 45

Leu Arg Leu Ser Glu Pro Leu Val Thr Ser Asn Gly Met Leu Ala Leu 50 55 60

Lys Met Gly Asn Gly Leu Ser Leu Asp Glu Ala Gly Asn Leu Thr Ser 65 70 75 80

Gln Asn Val Thr Thr Val Ser Pro Pro Leu Lys Lys Thr Lys Ser Asn 85 90 95

Ile Asn Leu Glu Ile Ser Ala Pro Leu Thr Val Thr Ser Glu Ala Leu 100 105 110

Thr Val Ala Ala Ala Ala Pro Leu Met Val Ala Gly Asn Thr Leu Thr 115 120 125

Met Gln Ser Gln Ala Pro Leu Thr Val His Asp Ser Lys Leu Ser Ile 130 135 140

Ala Thr Gln Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu Gln 145 150 155 160

Thr	Ser	Gly	Pro	Leu	Thr	Thr	Thr	Asp	Ser	Ser	Thr	Leu	Thr	Ile	Thr
				165					170					175	

Ala Ser Pro Pro Leu Thr Thr Ala Thr Gly Ser Leu Gly Ile Asp Leu 180 185 190

Lys Glu Pro Ile Tyr Thr Gln Asn Gly Lys Leu Gly Leu Lys Tyr Gly
195 200 205

Ala Pro Leu His Val Thr Asp Asp Leu Asn Thr Leu Thr Val Ala Thr 210 215 220

Gly Pro Gly Val Thr Ile Asn Asn Thr Ser Leu Gln Thr Lys Val Thr 225 230 235 240

Gly Ala Leu Gly Phe Asp Ser Gln Gly Asn Met Gln Leu Asn Val Ala
245
250
255

Gly Gly Leu Arg Ile Asp Ser Gln Asn Arg Arg Leu Ile Leu Asp Val 260 265 270

Ser Tyr Pro Phe Asp Ala Gln Asn Gln Leu Asn Leu Arg Leu Gly Gln 275 280 285

Gly Pro Leu Phe Ile Asn Ser Ala His Asn Leu Asp Ile Asn Tyr Asn 290 295 300 Lys Gly Leu Tyr Leu Phe Thr Ala Ser Asn Asn Ser Lys Lys Leu Glu 305 310 315 320

Val Asn Leu Ser Thr Ala Lys Gly Leu Met Phe Asp Ala Thr Ala Ile
325
330
335

Ala Ile Asn Ala Gly Asp Gly Leu Glu Phe Gly Ser Pro Asn Ala Pro 340 345 350

Asn Thr Asn Pro Leu Lys Thr Lys Ile Gly His Gly Leu Glu Phe Asp 355 360 365

Ser Asn Lys Ala Met Val Pro Lys Leu Gly Thr Gly Leu Ser Phe Asp 370 375 380

Ser Thr Gly Ala Ile Thr Val Gly Asn Lys Asn Asn Asp Lys Leu Thr 385 390 395 400

Leu Trp Thr Thr Pro Ala Pro Ser Pro Asn Cys Arg Leu Asn Ala Glu
405 410 415

Lys Asp Ala Lys Leu Thr Leu Val Leu Thr Lys Cys Gly Ser Gln Ile
420 425 430

Leu Ala Thr Val Ser Val Leu Ala Val Lys Gly Ser Leu Ala Pro Ile

Ser Gly Thr Val Gln Ser Ala His Leu Ile Ile Arg Phe Asp Glu Asn 450 455 460

Gly Val Leu Leu Asn Asn Ser Phe Leu Asp Pro Glu Tyr Trp Asn Phe
465 470 475 480

Arg Asn Gly Asp Leu Thr Glu Gly Thr Ala Tyr Thr Asn Ala Val Gly
485
490
495

Phe Met Pro Asn Leu Ser Ala Tyr Pro Lys Ser His Gly Lys Thr Ala 500 505 510

Lys Ser Asn Ile Val Ser Gln Val Tyr Leu Asn Gly Asp Lys Thr Lys
515 520 525

Pro Val Thr Leu Thr Ile Thr Leu Asn Gly Thr Gln Glú Thr Gly Asp 530 535 540

Thr Thr Pro Ser Ala Tyr Ser Met Ser Phe Ser Trp Asp Trp Ser Gly 545 550 555 560

His Asn Tyr Ile Asn Glu Ile Phe Ala Thr Ser Ser Tyr Thr Phe Ser 565 570 575

Tyr	Ile	Ala	Gln	Glu	Pro	Ser	Ala	Ser	Ala	Ser	Ala	Ser	Ala	Pro	Gly
			580					585					590		

Lys Lys Lys Lys Lys Lys Lys Gly Ser Tyr Ser Met Glu His Phe 610 615 620

Arg Trp Gly Lys Pro Val

625 630

<210> 37

<211> 630

<212> PRT

<213> Artificial Sequence

<220>

<223> a modified fiber protein encoded in pWE6.7R-F/gsK21MSHa

<400> 37

Met Lys Arg Ala Arg Pro Ser Glu Asp Thr Phe Asn Pro Val Tyr Pro 1 5 10 15

Tyr Asp Thr Glu Thr Gly Pro Pro Thr Val Pro Phe Leu Thr Pro Pro

Phe Val Ser Pro Asn Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser

Leu Arg Leu Ser Glu Pro Leu Val Thr Ser Asn Gly Met Leu Ala Leu
50 55 60

Lys Met Gly Asn Gly Leu Ser Leu Asp Glu Ala Gly Asn Leu Thr Ser 65 70 75 80

Gln Asn Val Thr Thr Val Ser Pro Pro Leu Lys Lys Thr Lys Ser Asn 85 90 95

Ile Asn Leu Glu Ile Ser Ala Pro Leu Thr Val Thr Ser Glu Ala Leu 100 105 110

Thr Val Ala Ala Ala Pro Leu Met Val Ala Gly Asn Thr Leu Thr
115 120 125

Met Gln Ser Gln Ala Pro Leu Thr Val His Asp Ser Lys Leu Ser Ile 130 135 140

Ala Thr Gln Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu Gln 145 150 155 160 Thr Ser Gly Pro Leu Thr Thr Asp Ser Ser Thr Leu Thr Ile Thr
165 170 175

Ala Ser Pro Pro Leu Thr Thr Ala Thr Gly Ser Leu Gly Ile Asp Leu 180 185 190

Lys Glu Pro Ile Tyr Thr Gln Asn Gly Lys Leu Gly Leu Lys Tyr Gly
195 200 205

Ala Pro Leu His Val Thr Asp Asp Leu Asn Thr Leu Thr Val Ala Thr 210 215 220

Gly Pro Gly Val Thr Ile Asn Asn Thr Ser Leu Gln Thr Lys Val Thr
225 230 235 240

Gly Ala Leu Gly Phe Asp Ser Gln Gly Asn Met Gln Leu Asn Val Ala 245 250 255

Gly Gly Leu Arg Ile Asp Ser Gln Asn Arg Arg Leu Ile Leu Asp Val 260 265 270

Ser Tyr Pro Phe Asp Ala Gln Asn Gln Leu Asn Leu Arg Leu Gly Gln
275
280
285

Gly Pro Leu Phe Ile Asn Ser Ala His Asn Leu Asp Ile Asn Tyr Asn 290 295 300 Lys Gly Leu Tyr Leu Phe Thr Ala Ser Asn Asn Ser Lys Lys Leu Glu 305 310 315 320

Val Asn Leu Ser Thr Ala Lys Gly Leu Met Phe Asp Ala Thr Ala Ile 325 330 335

Ala Ile Asn Ala Gly Asp Gly Leu Glu Phe Gly Ser Pro Asn Ala Pro 340 345 350

Asn Thr Asn Pro Leu Lys Thr Lys Ile Gly His Gly Leu Glu Phe Asp 355 360 365

Ser Asn Lys Ala Met Val Pro Lys Leu Gly Thr Gly Leu Ser Phe Asp 370 375 380

Ser Thr Gly Ala Ile Thr Val Gly Asn Lys Asn Asn Asp Lys Leu Thr 385 390 395 400

Leu Trp Thr Thr Pro Ala Pro Ser Pro Asn Cys Arg Leu Asn Ala Glu
405 410 415

Lys Asp Ala Lys Leu Thr Leu Val Leu Thr Lys Cys Gly Ser Gln Ile
420 425 430

Leu Ala Thr Val Ser Val Leu Ala Val Lys Gly Ser Leu Ala Pro Ile

435 440 445

Ser Gly Thr Val Gln Ser Ala His Leu Ile Ile Arg Phe Asp Glu Asn 450 455 460

Gly Val Leu Leu Asn Asn Ser Phe Leu Asp Pro Glu Tyr Trp Asn Phe 465 470 475 480

Arg Asn Gly Asp Leu Thr Glu Gly Thr Ala Tyr Thr Asn Ala Val Gly
485 490 495

Phe Met Pro Asn Leu Ser Ala Tyr Pro Lys Ser His Gly Lys Thr Ala
500 505 510

Lys Ser Asn Ile Val Ser Gln Val Tyr Leu Asn Gly Asp Lys Thr Lys
515 520 525

Pro Val Thr Leu Thr Ile Thr Leu Asn Gly Thr Gln Glu Thr Gly Asp 530 535 540

Thr Thr Pro Ser Ala Tyr Ser Met Ser Phe Ser Trp Asp Trp Ser Gly
545 550 555 560

His Asn Tyr Ile Asn Glu Ile Phe Ala Thr Ser Ser Tyr Thr Phe Ser 565 570 575

Tyr Ile Ala Gln Glu Gly Ser Gl

Lys Lys Lys Lys Lys Lys Lys Gly Ser Tyr Ser Met Glu His Phe 610 615 620

Arg Trp Gly Lys Pro Val

625

630

<210> 38

<211> 640

<212> PRT

<213> Artificial Sequence

<220>

<223> a modified fiber protein encoded in pWE6. 7R-F/asK21MSHb

<400> 38

Met Lys Arg Ala Arg Pro Ser Glu Asp Thr Phe Asn Pro Val Tyr Pro 1 5 10 15

Tyr Asp Thr Glu Thr Gly Pro Pro Thr Val Pro Phe Leu Thr Pro Pro

Phe Val Ser Pro Asn Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser 35 40 45

Leu Arg Leu Ser Glu Pro Leu Val Thr Ser Asn Gly Met Leu Ala Leu 50 55 60

Lys Met Gly Asn Gly Leu Ser Leu Asp Glu Ala Gly Asn Leu Thr Ser 65 70 75 80

Gln Asn Val Thr Thr Val Ser Pro Pro Leu Lys Lys Thr Lys Ser Asn 85 90 95

Ile Asn Leu Glu Ile Ser Ala Pro Leu Thr Val Thr Ser Glu Ala Leu 100 105 110

Thr Val Ala Ala Ala Pro Leu Met Val Ala Gly Asn Thr Leu Thr 115 120 125

Met Gln Ser Gln Ala Pro Leu Thr Val His Asp Ser Lys Leu Ser Ile 130 135 140

Ala Thr Gln Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu Gln 145 150 155 160

Thr	Ser	Gly	Pro	Leu	Thr	Thr	Thr	Asp	Ser	Ser	Thr	Leu	Thr	Ile	Thr
				165					170					175	

Ala Ser Pro Pro Leu Thr Thr Ala Thr Gly Ser Leu Gly Ile Asp Leu 180 185 190

Lys Glu Pro Ile Tyr Thr Gln Asn Gly Lys Leu Gly Leu Lys Tyr Gly
195 200 205

Ala Pro Leu His Val Thr Asp Asp Leu Asn Thr Leu Thr Val Ala Thr 210 215 220

Gly Pro Gly Val Thr Ile Asn Asn Thr Ser Leu Gln Thr Lys Val Thr 225 230 235 240

Gly Ala Leu Gly Phe Asp Ser Gln Gly Asn Met Gln Leu Asn Val Ala
245
250
255

Gly Gly Leu Arg Ile Asp Ser Gln Asn Arg Arg Leu Ile Leu Asp Val 260 265 270

Ser Tyr Pro Phe Asp Ala Gln Asn Gln Leu Asn Leu Arg Leu Gly Gln
275
280
285

Gly Pro Leu Phe Ile Asn Ser Ala His Asn Leu Asp Ile Asn Tyr Asn 290 295 300

Lys	Gly	Leu	Tyr	Leu	Phe	Thr	Ala	Ser	Asn	Asn	Ser	Lys	Lys	Leu	Glu
305	٠				310					315					320

Val Asn Leu Ser Thr Ala Lys Gly Leu Met Phe Asp Ala Thr Ala Ile 325 330 335

Ala Ile Asn Ala Gly Asp Gly Leu Glu Phe Gly Ser Pro Asn Ala Pro 340 345 350

Asn Thr Asn Pro Leu Lys Thr Lys Ile Gly His Gly Leu Glu Phe Asp 355 360 365

Ser Asn Lys Ala Met Val Pro Lys Leu Gly Thr Gly Leu Ser Phe Asp 370 375 380

Ser Thr Gly Ala Ile Thr Val Gly Asn Lys Asn Asn Asp Lys Leu Thr 385 390 395 400

Leu Trp Thr Thr Pro Ala Pro Ser Pro Asn Cys Arg Leu Asn Ala Glu
405 410 415

Lys Asp Ala Lys Leu Thr Leu Val Leu Thr Lys Cys Gly Ser Gln Ile
420 425 430

Leu Ala Thr Val Ser Val Leu Ala Val Lys Gly Ser Leu Ala Pro Ile

hik

435 440 445

Ser Gly Thr Val Gln Ser Ala His Leu Ile Ile Arg Phe Asp Glu Asn 450 455 460

Gly Val Leu Leu Asn Asn Ser Phe Leu Asp Pro Glu Tyr Trp Asn Phe 465 470 475 480

Arg Asn Gly Asp Leu Thr Glu Gly Thr Ala Tyr Thr Asn Ala Val Gly
485 490 495

Phe Met Pro Asn Leu Ser Ala Tyr Pro Lys Ser His Gly Lys Thr Ala
500 505 510

Lys Ser Asn Ile Val Ser Gln Val Tyr Leu Asn Gly Asp Lys Thr Lys
515 520 525

Pro Val Thr Leu Thr Ile Thr Leu Asn Gly Thr Gln Glu Thr Gly Asp 530 535 540

Thr Thr Pro Ser Ala Tyr Ser Met Ser Phe Ser Trp Asp Trp Ser Gly
545 550 555 560

His Asn Tyr Ile Asn Glu Ile Phe Ala Thr Ser Ser Tyr Thr Phe Ser 565 570 575

Tyr Ile Ala Gln Glu Pro Ser Ala Ser Ala Ser Ala Ser Ala Pro Gly
580 585 590

Lys Lys Lys Lys Lys Lys Gly Ser Ala Glu Lys Lys Asp Glu
610 615 620

Gly Pro Tyr Arg Met Glu His Phe Arg Trp Gly Ser Pro Pro Lys Asp 625 630 635 640

<210> 39

<211> 640

<212> PRT

<213> Artificial Sequence

<220>

<223> a modified fiber protein encoded in pWE6.7R-F/gsK21MSHb

<400> 39

Met Lys Arg Ala Arg Pro Ser Glu Asp Thr Phe Asn Pro Val Tyr Pro

1 5 10 15

Tyr Asp Thr Glu Thr Gly Pro Pro Thr Val Pro Phe Leu Thr Pro Pro

20

21.3 214
i.)
1.1
p de
12
jari.
L/I
h .5
#
CJ
ru
edita isa
C)
-

Phe	Val	Ser	Pro	Asn	Gly	Phe	Gln	Glu	Ser	Pro	Pro	Gly	Val	Leu	Ser
		35					40					45			

Leu Arg Leu Ser Glu Pro Leu Val Thr Ser Asn Gly Met Leu Ala Leu 50 55 60

Lys Met Gly Asn Gly Leu Ser Leu Asp Glu Ala Gly Asn Leu Thr Ser
65 70 75 80

Gln Asn Val Thr Thr Val Ser Pro Pro Leu Lys Lys Thr Lys Ser Asn 85 90 95

Ile Asn Leu Glu Ile Ser Ala Pro Leu Thr Val Thr Ser Glu Ala Leu 100 105 110

Thr Val Ala Ala Ala Ala Pro Leu Met Val Ala Gly Asn Thr Leu Thr
115 120 125

Met Gln Ser Gln Ala Pro Leu Thr Val His Asp Ser Lys Leu Ser Ile 130 135 140

Ala Thr Gln Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu Gln 145 150 155 160

Thr	Ser	Gly	Pro	Leu	Thr	Thr	Thr	Asp	Ser	Ser	Thr	Leu	Thr	Ile	Thr
				165					170					175	

Ala Ser Pro Pro Leu Thr Thr Ala Thr Gly Ser Leu Gly Ile Asp Leu 180 185 190

Lys Glu Pro Ile Tyr Thr Gln Asn Gly Lys Leu Gly Leu Lys Tyr Gly
195 200 205

Ala Pro Leu His Val Thr Asp Asp Leu Asn Thr Leu Thr Val Ala Thr
210 215 220

Gly Pro Gly Val Thr Ile Asn Asn Thr Ser Leu Gln Thr Lys Val Thr
225 230 235 240

Gly Ala Leu Gly Phe Asp Ser Gln Gly Asn Met Gln Leu Asn Val Ala 245 250 255

Gly Gly Leu Arg Ile Asp Ser Gln Asn Arg Arg Leu Ile Leu Asp Val
260 265 270

Ser Tyr Pro Phe Asp Ala Gln Asn Gln Leu Asn Leu Arg Leu Gly Gln
275
280
285

Gly Pro Leu Phe Ile Asn Ser Ala His Asn Leu Asp Ile Asn Tyr Asn 290 295 300

Lys	Gly	Leu	Tyr	Leu	Phe	Thr	Ala	Ser	Asn	Asn	Ser	Lys	Lys	Leu	Glu
305					310					315					320

Val Asn Leu Ser Thr Ala Lys Gly Leu Met Phe Asp Ala Thr Ala Ile 325 330 335

Ala Ile Asn Ala Gly Asp Gly Leu Glu Phe Gly Ser Pro Asn Ala Pro 340 345 350

Asn Thr Asn Pro Leu Lys Thr Lys Ile Gly His Gly Leu Glu Phe Asp 355 360 365

Ser Asn Lys Ala Met Val Pro Lys Leu Gly Thr Gly Leu Ser Phe Asp 370 375 380

Ser Thr Gly Ala Ile Thr Val Gly Asn Lys Asn Asn Asp Lys Leu Thr 385 390 395 400

Leu Trp Thr Thr Pro Ala Pro Ser Pro Asn Cys Arg Leu Asn Ala Glu
405 410 415

Lys Asp Ala Lys Leu Thr Leu Val Leu Thr Lys Cys Gly Ser Gln Ile
420 425 430

Leu Ala Thr Val Ser Val Leu Ala Val Lys Gly Ser Leu Ala Pro Ile

Ser	Gly	Thr	Val	Gln	Ser	Ala	His	Leu	Ile	Ile	Arg	Phe	Asp	Glu	Asn
	450					455				•	460				

Gly Val Leu Leu Asn Asn Ser Phe Leu Asp Pro Glu Tyr Trp Asn Phe 465 470 475 480

Arg Asn Gly Asp Leu Thr Glu Gly Thr Ala Tyr Thr Asn Ala Val Gly
485 490 495

Phe Met Pro Asn Leu Ser Ala Tyr Pro Lys Ser His Gly Lys Thr Ala
500 505 510

Lys Ser Asn Ile Val Ser Gln Val Tyr Leu Asn Gly Asp Lys Thr Lys
515 520 525

Pro Val Thr Leu Thr Ile Thr Leu Asn Gly Thr Gln Glu Thr Gly Asp 530 535 540

Thr Thr Pro Ser Ala Tyr Ser Met Ser Phe Ser Trp Asp Trp Ser Gly 545 550 555 560

His Asn Tyr Ile Asn Glu Ile Phe Ala Thr Ser Ser Tyr Thr Phe Ser 565 570 575



Tyr Ile Ala Gln Glu Gly Ser Gl

Lys Lys Lys Lys Lys Lys Lys Gly Ser Ala Glu Lys Lys Asp Glu 610 615 620

Gly Pro Tyr Arg Met Glu His Phe Arg Trp Gly Ser Pro Pro Lys Asp 625 630 630 635